A Hidden Square-Root Boundary Between Growth Rate and Biomass Yield

Wilson W. Wong, Linh M. Tran, James C. Liao
Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, California 90095, Telephone: +1-310-825-1656; Fax: +1-310-206-4107; email: liaoj@ucla.edu

Received 10 January 2008; revision received 18 May 2008; accepted 7 July 2008
Published online 15 July 2008 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.22046

ABSTRACT: Although the theoretical value of biomass yield can be calculated from metabolic network stoichiometry, the growth rate is difficult to predict. Since the rate and yield can vary independently, no simple relationship has been discovered between these two variables. In this work, we analyzed the well-accepted enzyme kinetics and uncovered a hidden boundary for growth rate, which is determined by the square-root of three physiological parameters: biomass yield, the substrate turnover number, and the maximum synthesis rate of the turnover enzyme. Cells cannot grow faster than the square-root of the product of these parameters. This analysis is supported by experimental data and involves essentially no assumptions except (i) the cell is not undergoing a downshift transition, (ii) substrate uptake enzyme activity is proportional to its copy number. This simple boundary (not correlation) has escaped notice for many decades and suggests that the yield calculation does not predict the growth rate, but gives an upper limit for the growth rate. The relationship also explains how growth rate is affected by the yield and sheds lights on strain design for product formation.

© 2008 Wiley Periodicals, Inc.
KEYWORDS: metabolic network; microbial growth; theoretical evolution boundary

Introduction

Growth rate and biomass yield are important physiological measurements used in microbiology, biochemical engineering, and industrial biotechnology. They are particularly significant in large-scale production of biofuels and commodity chemicals where efficiency and rate are crucial. They are also important in understanding physiology and evolution in microbiology. While the specific growth rate of a microorganism is a kinetic property which reflects the time efficiency, biomass yield (biomass produced per substrate consumed) is a dimensionless quantity reflecting the resource utilization efficiency. By definition, these two variables are related by the specific substrate consumption rate: the specific growth rate divided by the specific substrate consumption rate is the biomass yield. Since the specific substrate consumption rate is a strong function of intracellular states and extracellular conditions, no obvious relationship between growth rate and the yield exists. However, these two quantities are often implicitly assumed proportional to each other in strain design, or inversely proportional to each other when discussing the trade-off between rate and efficiency in evolution (Pfeiffer et al., 2001). Numerous models exist for growth rate (Button, 1985) and yield (Edwards and Palsson, 2000; Liao et al., 1996) separately, but none explicitly discussed their relationship.

Organisms such as Saccharomyces cerevisiae and Escherichia coli use more wasteful fermentative pathways even under aerobic conditions (Farmer and Liao, 1997; Picon et al., 2005; Postma et al., 1989; MacLean and Gudelj, 2006), which result in a lower yield while increasing growth rate. Many plausible explanations exist, including the limitation of TCA cycle, the optimal ATP production rate, and a trade-off between the rate and yield (Kreft, 2004; Pfeiffer et al., 2001), which implies that growth rate might decrease when the biomass yield increases by a more efficient metabolic network design. Other theories indirectly related to the rate and the yield involve protein synthesis that optimizes the trade-off between the cost and benefit (Dekel and Alon, 2005; Young and Ramkrishna, 2007). In general, the relationship between the growth rate and yield has been obscure, and the relationship may be explored from multiple angles emphasizing different aspects of physiology. Nevertheless, this relationship is fundamental to evolution and microbial strain design and remains to be further illuminated. Does efficient resource utilization necessarily
leads to slower or faster growth? To what extent does a metabolic engineering strategy aiming to improve product yield impact growth rate?

An obvious but impractical way to answer these questions is to obtain the complete kinetics and interacting networks of all molecules in the cell. One then constructs a detailed model describing each reaction and interaction. Given the practical difficulty in this brute-force approach, we seek constraints that may shed light on this fundamental relationship. Here we use a set of general, but unspecified kinetic equations with measurable parameters as a starting point of analysis. From this general model, we derived an inequality that serves as the upper bound for growth rate in terms of measurable physiological quantities. We present in the following the theoretical arguments and experimental supports based on growth rate and biomass yield measurements. Note that our derivation is based on mass balance equations and common fermentation kinetics without assumption of rate limiting steps, and that the result is not a correlation, but an upper boundary determined by mass balance and kinetics.

**Materials and Methods**

**Bacterial Strains and Plasmids**

*E. coli* mutants were obtained either from our lab’s collection, from a transposon library (Nichols et al., 1998; Singer et al., 1989), or from the Keio Collection (Baba et al., 2006). Some of the knockout mutants were obtain using the method by Datsenko and Wanner (2000). All experiments involving lactose were done using MG1655. MG1655Δcrr and MG1655ΔptsG were obtained from P1 transduction using the Keio Collection as the donor.

**Culture Conditions**

*E. coli* MG1655 and various other mutants was grown in M9 minimal media supplemented with 0.1% of carbon source (lactose or glucose), 1 mM MgSO4, 1 μg/mL vitamin B1, 100 μM CaCl2, or when needed, various amino acids and incubated in a 37°C water bath shaking at 250 rpm. With the initial OD of 0.05, the cultures were allowed to double once before various measurements were taken to ensure that the cells are grown out of the stationary phase.

At each time point, samples were collected for OD measurement, or lactose determination. The sampling interval is determined by the growth rate. For wild-type strains or mutants with a growth rate similar to the wild-type, the typical sampling interval was between 15 and 30 min. For slower growing strains, the typical interval was between 30 and 90 min. At least five time points were taken per culture. The samples were grown to a final OD of 0.3.

For lactose determination, 700 μL of culture was filtered immediately with a syringe filter (0.22 μm) to prevent further consumption of lactose by the cells. OD was measured by using a spectrophotometer at 600 nm (Beckman DU520). In the lactose concentration measurement, filtered supernatant was injected into HPLC with auto-sampler (Agilent 1100) equipped with a reflective index detector and an Aminex HPX87 column (0.5 mM H2SO4, 0.6 mL/min). Concentration was determined was extrapolated from standard curves. The biomass yield is determined by the linear regression coefficient of OD against the concentration of extracellular lactose. OD is converted to dry cell weight by the conversion factor, 1 OD = 0.33 gdcw/mL.

**Anaerobic Experiments**

Some of the wild-type rate–yield relationship experiments were performed anaerobically. Cultures were grown in media contained in a sealed, airtight bottle where the air inside the bottle was replaced with argon. The growth experiment was performed under normal conditions in a water bath shaker. Samples were removed with needles and syringes. OD and biomass yield measurements were performed as described earlier.

**Theoretical Derivation**

**Theoretical Analysis Identifies a Square-Root Boundary**

We start from a commonly accepted kinetic description of substrate (e.g., carbon source, nitrogen source, and other nutrients) consumption, which assumes that enzyme activity is proportional to its concentration multiplied by an unspecified saturation kinetic function.

\[
\frac{dx_1}{dt} = -y_1 k_{cat} g_1
\]

where \(x_1\) [substrate mass/culture volume] is the concentration of the substrate in the medium, \(y_1\) [transporter/culture volume] is the total concentration of the substrate transporter (or initial turnover enzyme) in the culture [\(y_1 = (\text{transporter per cell mass}) \times (\text{cell mass per volume})\)], \(k_{cat}\) [substrate turnover/time/transporter] is the turnover number for the substrate uptake transporter and \(g_1\) [dimensionless] is an unspecified kinetic function of substrate transport whose maximal is unity. \(g_1\) can be a function of the substrate and other metabolites, through which the cell controls the substrate uptake rate. According to this equation, the maximum uptake rate under this condition is \(y_1 k_{cat}\).

The synthesis and degradation of each of the transporters (or the initial turnover enzyme) is accounted for by:

\[
\frac{dy_1}{dt} = BR_m f_1 - k_{deg} y_1
\]
Here, \( B \) [cell mass/culture volume] is the total mass of cells per culture volume, \( R_m \) [transporter/time/cell mass] is the maximum synthesis rate of the transporter per cell mass. \( f_t \) [dimensionless] represents an unspecified kinetic function of the transporter synthesis rate scaled to unity, through which the cell controls the synthesis of \( y_1 \). \( k_{\text{deg}} \) [1/time] is the first-order kinetic constant for the degradation of the transporter. The degradation of \( y_1 \) can be zero or more complicated functions. Since the kinetic form of protein degradation will not change subsequent analysis, for simplicity we use first-order kinetics. It is worth noting that our definition of \( y_1 \) is unconventional, because of its convenience in theoretical derivation. Note that we did not assume the substrate uptake is the rate limiting step.

From Equations (1) and (2), we have:

\[
k_{\text{cat}} = \left( \frac{-dx_1}{dt} \right) \frac{1}{y_1g_1} \quad (3)
\]

\[
R_m = \frac{1}{f_t} \left( \frac{dy_1}{dt} + k_{\text{deg}}y_1 \right) \quad (4)
\]

To describe the cell growth rate completely, the rest of the metabolic and regulatory system needs to be accounted for. However, regardless of the kinetics of the rest of the system, Equations (1) and (2) are valid, and biomass can be expressed as a function of any substrate consumed with an easily measurable yield coefficient: \( (B-B_0) = Y_B (x_1-x_{1,0}) \), where \( Y_B \) is the biomass yield with respect to \( x_1 \) and the subscript 0 represents the initial concentrations. Since \( x_1 \) is the only carbon source, the cell density can be expressed as a function of \( x_1 \): \( B = f(x_1) \). Therefore:

\[
\frac{dB}{dt} = \frac{dB}{dx_1} \frac{dx_1}{dt} \quad \Leftrightarrow \quad \frac{dx_1}{dt} = \frac{dB}{dt} \left( \frac{dB}{dx_1} \right)^{-1} \quad (5)
\]

We define that \( y_1 = \tilde{y}_1 B \), where \( \tilde{y}_1 \) is amount of the transporters per cell. We have:

\[
\frac{dy_1}{dt} = \tilde{y}_1 \frac{dB}{dt} + B \frac{dy_1}{dt} \quad (6)
\]

Multiplying (3) and (4) together, then substituting (5) and (6) into the product, we obtain the following equation:

\[
k_{\text{cat}}R_m = \left( -\frac{dB}{dx_1} \right)^{-1} \frac{1}{f_tg_1} \left( \frac{1}{dB} \frac{dB}{dt} \right)^2 + \frac{1}{y_1} \frac{dB}{dt} \left( \frac{dB}{x_1} \right) + k_{\text{deg}} \left( \frac{1}{B} \frac{dB}{dt} \right) \quad (7)
\]

The biomass yield, \( Y_B \), is defined as

\[
Y_B \equiv -\frac{dB}{dx_1} \quad (8)
\]

and the specific growth rate, \( \mu \), is defined as:

\[
\mu = \frac{1}{B} \frac{dB}{dt} \quad (9)
\]

After substituting these definitions into Equation (7) we have,

\[
\mu^2 + \left( \frac{d \ln y_1}{dt} + k_{\text{deg}} \right) \mu - Y_B R_m k_{\text{cat}} f_1 g_1 = 0 \quad (10)
\]

Therefore

\[
\mu = \frac{-\left( \frac{d \ln y_1}{dt} + k_{\text{deg}} \right) + \sqrt{\left( \frac{d \ln y_1}{dt} + k_{\text{deg}} \right)^2 + 4Y_B R_m k_{\text{cat}} f_1 g_1}}{2} \quad (11)
\]

In addition to the biomass yield, the protein synthesis rate, and the substrate uptake turnover rate (i.e., \( R_m, f_t \) and \( k_{\text{cat}}, g_1 \)), the growth rate also depends on the protein degradation rate constant, \( k_{\text{deg}} \).

Let us define \( S \equiv (d \ln y_1/dt) = (1/y_1)(dy_1/dt) \), which can be either: (i) equal to 0 in balanced growth, (ii) greater than 0 in upshifted growth (i.e., the protein synthesis rate is greater than the degradation rate and the dilution rate), or (iii) less than 0 in downshifted growth (i.e., the protein synthesis rate is less than the degradation and the dilution rates).

Interestingly, For \( S \geq 0 \) (balanced growth or upshift), Equation (11) is maximized as \( g_1 = f_1 = 1, k_{\text{deg}} = 0 \) and \( S = 0 \). Therefore, the following inequality emerges:

\[
\mu \leq \sqrt{Y_B R_m k_{\text{cat}}} \quad (12)
\]

For \( S < 0 \) (downshift) and \( k_{\text{deg}} = 0 \), it is easy to show that:

\[
\mu > \sqrt{Y_B R_m k_{\text{cat}} g_1 f_1} \quad (13)
\]

Therefore, it is possible that \( \mu \) is greater than \( \sqrt{Y_B R_m k_{\text{cat}}} \). In most cases, however, \( g_1 \) and \( f_1 \) will be sufficiently reduced during a downshift, and \( \mu \) may still be less than \( \sqrt{Y_B R_m k_{\text{cat}}} \).

In the case of \( k_{\text{deg}} = 0, g_1 = 1 \) and \( f_1 = 1 \), then \( \mu = \sqrt{Y_B R_m k_{\text{cat}}} \). Note that \( Y_B \) and \( \mu \) do not have to be constant and that the above equation can apply to time-dependent or time averaged measurements. The applicability of this relationship is discussed below.

**Downstream Factors Limit Cell Growth**

The derivation above may give a false impression that substrate uptake is implicitly assumed to be the limiting step for cell growth. No such assumption was made. In cases where intracellular metabolism limits cell growth, the above derivation still applies as long as Equations (1) and (2) are valid. Since the cell cannot indefinitely accumulate
metabolites internally, the cell will either reduce the substrate uptake rate, secrete metabolites that could not be further digested, or store a product intracellularly.

If the substrate uptake flux is to be reduced, this action must be achieved through either inhibiting transporter activity ($g_i < 1$), synthesizing less transporter ($f_i < 1$), or degrading existing transporter ($k_{deg} > 0$). Under these conditions, Equation (12) reduces to $\mu < \sqrt{Y_B R_m k_{cat}}$. Under these conditions, Equation (12) reduces to $\mu < \sqrt{Y_B R_m k_{cat}}$. Under these conditions, Equation (12) reduces to $\mu < \sqrt{Y_B R_m k_{cat}}$. Under these conditions, Equation (12) reduces to $\mu < \sqrt{Y_B R_m k_{cat}}$.

If metabolites are secreted, then $Y_B$ decreases, and Equation (12) still holds. If a compound is stored as an intracellular product (e.g., glycogen), $R_m$ will be reduced because of the higher cell mass per cell. However, Equation (12) still applies after the cell reaches a new steady state. The equality part of Equation (12) may still apply to these two cases.

**Multiple Substrates Are Consumed Simultaneously Through Different Transporters**

Carbon, nitrogen, and other nutrients are simultaneously consumed to support cell growth. Each nutrient can be described separately using the general form of Equations (1) and (2). Typically, the nutrients are transported via protein-mediated activity or facilitated transporters that exhibit a saturation behavior and thus the above equations hold. The coordination of consumption rates for different nutrients is exerted through each of the unspecified kinetic function $g_i$ and $f_i$. $Y_B$ can be defined for each nutrient of interest, and thus Equation (12) applies to each of the nutrients individually. If multiple carbon sources (e.g., amino acids and glucose) are simultaneously consumed, the biomass yield defined based on each carbon source will increase. The biomass yield defined based on each nutrient is related to each other through the overall cellular nutritional requirements and mass balance. Nevertheless, Equation (12) applies to each nutrient.

**The Substrate Is Freely Diffusible Through the Membrane Without a Transporter**

For substrates that diffuse through the lipid bilayer (e.g., CO$_2$ and O$_2$) without a clear maximum activity, then $f_i$ should be defined as the initial enzyme that consumes the substrate (e.g., RuBisco in CO$_2$-fixing organisms).

**Multiple Substrates Are Transported Through the Same Transporter Simultaneously**

As described above, each substrate can be considered separately and the $Y_B$ defined for each substrate will increase. The kinetic function $g_i$ needs to reflect the competition between the two substrates. In the presence of competing substrate, $g_i$ will be less than unity, but Equation (12) still holds.

**Multiple Transporters Are Used for a Single Substrate**

The consumption of carbon source can be described with

$$\frac{dx_i}{dt} = -\sum_{i=1}^{n} y_i k_{cat,i} g_i$$

(14)

where $n$ is the number of transporters. The synthesis and degradation of the transporter $i$ is accounted for by:

$$\frac{dy_i}{dt} = BR_m y_i - k_{deg,i} y_i$$

(15)

From the biomass yield definition, we have:

$$\frac{dB}{dt} = -Y_B \frac{dx_i}{dt} = Y_B \sum_{i=1}^{n} y_i k_{cat,i} g_i$$

(16)

Similar to the previous section, we define that $y_i = \tilde{y}_i B$, where $\tilde{y}_i$ is amount of the transporters per cell, then:

$$\frac{dy_i}{dt} = \tilde{y}_i \frac{dB}{dt} + B \frac{dy_i}{dt}$$

(17)

Equating Equations (15) and (17), then substituting the definition of the growth rate, we get:

$$\mu + \frac{1}{\tilde{y}_i} \frac{dy_i}{dt} = \frac{R_m f_i}{\tilde{y}_i} - k_{deg,i}$$

(18)

or

$$\tilde{y}_i = \frac{R_m f_i}{\mu + k_{deg,i} + \frac{d \ln \tilde{y}_i}{dt}}$$

(19)

Substituting Equation (19) into Equation (16), we can describe the growth rate as

$$\mu \equiv \frac{1}{B} \frac{dB}{dt} = Y_B \sum_{i=1}^{n} \frac{R_m f_i k_{cat,i} g_i}{\mu + k_{deg,i} + \frac{d \ln \tilde{y}_i}{dt}}$$

(20)

Analogous to the previous section, we define $S_i \equiv (d \ln \tilde{y}_i / dt)$. To examine the effect of $k_{deg,i}$ on the growth rate, we take the partial derivative with respect to $k_{deg,i}$ on both sides of Equation (20) implicitly, we get:

$$\frac{\partial \mu}{\partial k_{deg,i}} = -\frac{Y_B R_m f_i k_{cat,i} g_i}{(\mu + k_{deg,i} + S_i)^2} \left( \frac{\partial \mu}{\partial k_{deg,i}} + 1 \right)$$

(21)
Therefore, the partial derivative of growth rate with respect to \( k_{\text{deg},i} \) becomes:

\[
\frac{\partial \mu}{\partial k_{\text{deg},i}} = -C_i(1 + C_i)^{-1} < 0
\]  

where

\[
C = \frac{Y_B R_{m,i} k_{\text{cat},i} S_i}{\left( \mu + k_{\text{deg},i} + S_i \right)^2} > 0
\]

This equation means the growth rate decreases as increasing the degradation of transporters. The similar result is obtained for \( S_i \).

For \( S_i \geq 0 \) for all \( i = 1, 2, \ldots, n \) (balanced growth or upshift), Equation (20) is maximized as \( k_{\text{deg},i} = 0 \) and \( S_i = 0 \), for all \( i = 1, 2, \ldots, n \), which means the cell is at balanced growth, therefore, the following inequality relationship emerges:

\[
\mu \leq \sqrt{\frac{Y_B R_{m,i} k_{\text{cat},i} S_i}{\left( \mu + k_{\text{deg},i} + S_i \right)^2}}
\]  

\[
(23)
\]

**Transient Upshift of \( y_1 \) During Growth**

This situation occurs when the transporter synthesis per cell is induced because of a shift in nutritional or environmental conditions and has not reached the new balanced steady state yet. Under this condition, the uptake enzyme is not sufficient to support the new growth rate and thus \( \mu < \sqrt{Y_B R_{m} k_{\text{cat}}} \).

**Transient Downshift in \( y_1 \) During Growth**

This situation occurs when cells are undergoing a transition from a nutritional state with high transporter per cell to a low transporter state. In this case the transporter per cell is transiently greater than its new steady state because of the lower enzyme synthesis rate. If the cell does not excrete metabolites or limit the substrate uptake rate by a feedback mechanisms that reduces \( g_i \), \( \mu \) may be greater than \( \sqrt{Y_B R_{m} k_{\text{cat}}} \). Although unlikely, this situation is theoretically possible and is the only scenario that violates Equation (12) when Equations (1) and (2) hold.

**When \( g_1 = 1, f_1 = 1, \text{and } k_{\text{deg}} = 0 \) During Balanced Growth**

During balanced growth, where all components increase at the same rate, and if \( g_1 = 1, f_1 = 1, \text{and } k_{\text{deg}} = 0 \), then Equation (12) becomes \( \mu = \sqrt{Y_B R_{m} k_{\text{cat}}} \). This is the maximum growth rate the cell can grow under the given nutritional condition.

**Derivation of the Relationship for Product Formation Rate**

**Product Formation Yield**

In general, the cell produces some useful product \( P \) along with the waste product \( w \) (CO\textsubscript{2} for example). The overall reaction is therefore:

\[
X_1 \rightarrow B + P + w
\]

Based on the conservation of mass, we have:

\[
Y_B + Y_P + Y_w = 1
\]  

\[
(26)
\]

where \( Y_P \) is the product yield, which is the amount of product formed per gram of substrate consumed. \( Y_w \) is the amount of waste or by-products formed per gram of substrate consumed. From (26) the growth yield relates to the product yield as:

\[
Y_B = 1 - Y_P - Y_w
\]

Thus we have:

\[
\mu \leq \sqrt{(1 - Y_P - Y_w) R_{m} k_{\text{cat}}}
\]  

\[
(28)
\]

**Results and Discussion**

**Experimental Support for the Square-Root Boundary**

According to the above derivation and discussion, the growth rate is limited by Equation (12) provided that the cell is not going through a downshift transition. To test the prediction of Equation (12) experimentally, we chose a reference state as a normalization basis to provide a frame for comparison and to control for experimental variations. The reference state typically is a wild-type strain in unrestricted balanced growth. Under this condition, the equality in Equation (12) holds and can serve as a normalization basis.

Equation (12) then becomes

\[
\frac{\mu}{\mu_{\text{ref}}} \leq \sqrt{\frac{Y_B}{Y_{B,\text{ref}}} \frac{R_{m}}{R_{m,\text{ref}}} \frac{k_{\text{cat}}}{k_{\text{cat},\text{ref}}}}
\]

\[
(29)
\]

Since the measurements of \( R_{m} \) and \( k_{\text{cat}} \) for a large number of strains are time consuming, we chose conditions where \( R_{m}/R_{m,\text{ref}} \) and \( k_{\text{cat}}/k_{\text{cat},\text{ref}} \) are less than or equal to unity and measured only \( \mu \) and \( Y_B \) for a large number of strains. Equation (29) then becomes

\[
\frac{\mu}{\mu_{\text{ref}}} \leq \sqrt{\frac{Y_B}{Y_{B,\text{ref}}}}
\]

\[
(30)
\]
Genetic Variations

We tested 153 *E. coli* single-gene insertion or deletion mutants randomly selected from a transposon insertion library (Singer et al., 1989) and the Keio collection (Baba et al., 2006). Since the inactivation of genes is unlikely to cause an increase in $R_m$ or $k_{cat}$, we expect that Equation (30) will be satisfied, barring experimental error and the unlikely event that the mutation inactivates a repressor and causes an increase in $R_m$. These strains were grown in glucose, and the growth rate and biomass yield with respect to glucose were measured. The aerobic culture of MG1656 in glucose minimal medium was used as a reference state. The biomass yields formed a normal distribution with a mean close to the wild-type yield (Fig. 1A), suggesting that the biomass yield can be readily altered in both directions. On the other hand, the impact of these mutations on growth rate was mostly negative (Fig. 1B), indicating that growth rate is close to optimal for the wild-type strain. Combining these two data, we see that the growth rate data were mostly below the square-root of $Y_B$ (Fig. 1C), as predicted by Equation (30). A few data slightly above the predicted line are attributed to measurement noise. Interestingly, the strains having high yield does not necessarily grow fast (e.g., Δcrr, Δpgi). On the other hand, the strain that grew fastest (e.g., rpoS) did not have the highest yield. These results indicate that the yield calculation based on stoichiometry does not predict the growth rate, but gives a boundary for the growth rate.

To support our prediction further, we replotted the data from Fischer and Sauer (2005), which measured the growth rate and yield of various 127 knockout mutants of *Bacillus subtilis*. Figure 1D shows that the relative growth rate is mostly less than or equal to the square-root of the relative yield, suggesting that the relationship described by Equation (30) is not unique to *E. coli*.

Nutritional Variation

In addition, we test Equation (30) by nutritional variations using glucose as the major carbon source. The aerobic culture (MG1656) in glucose minimal medium was used as a reference state. Various amino acids were added to aerobic and anaerobic cultures containing glucose. The growth rate and the biomass yield with respect to glucose were measured. Interestingly, we found that the data obey the equality boundary of Equation (30) (Fig. 2A). Similarly, we used various sugars (mannitol, arabinose, and lactose) as the sole carbon source under aerobic and anaerobic condition. Again, a square-root relationship between the relative growth rate

---

**Figure 1.** The growth rate and biomass yield relationship. **A:** The biomass yield of *E. coli* mutants formed a normal distribution with a mean equals to the wild-type yield. **B:** The growth rate of the *E. coli* mutants, however, was mostly smaller than the wild-type strain. **C:** Growth rates and biomass yields of *E. coli* deletion mutants in glucose. The relative growth rates (denoted by subscript r) of *E. coli* mutants in glucose minimal media were smaller than the square-root of the relative yields. All growth rates and yields were compared to the wild-type in aerobic condition without AA. **D:** Growth rates and biomass yields of *B. subtilis* deletion mutants in glucose (Fischer and Sauer, 2005).
and the relative yield was observed (Fig. 2B). These data suggest that $R_{\text{app}}^m$ and $k_{\text{cat}}^{\text{app}}$ do not change much between the test condition and the reference state and that the nutritional and oxygen variations do not affect $g_1$, $f_1$, and $k_{\text{deg}}$.

**Implication in Metabolic Engineering**

An immediate application of the square-root relation is in strain design for metabolite production. Microorganisms are commonly engineered to produce products of interest such as biofuels (Atsumi et al., 2007, 2008; Bermejo et al., 1998; Hanai et al., 2007; Yomano et al., 1998), amino acids (Lutke-Eversloh and Stephanopoulos, 2007) 1,3-propandiol (Zhu et al., 2002), isoprenoids (Farmer and Liao, 2000; Martin et al., 2003). The diversion of carbon to these products reduces the cell mass yield, $Y_B$. According to Equation (29), without increasing $R_{\text{app}}^m$ or $k_{\text{cat}}^{\text{app}}$, the growth rate is bound to decrease (Theoretical Aspects) when product formation occurs during the growth phase. This is a phenomenon commonly observed. Therefore, to increase productivity during the growth phase, one should increase $R_{\text{app}}^m$ to compensate for the decrease in $Y_B$.

**Implication on Rate and Yield Trade-Off**

Since the wild-type *E. coli* follows the equality of Equation (12), this relationship may be used to explain the choice of less-efficient metabolic pathway to gain growth rate. It is recognized that protein overexpression decreases the biomass yield because of the energy cost due to additional protein synthesis at the expense of synthesizing others. Thus, the biomass yield, $Y_B$, is negatively correlated with $R_m$ at high $R_m$. At a low $R_m$, $Y_B$ remains relative constant because the metabolic capacity of the cell can digest the substrate intake efficiently. A hypothetical example of such a relationship is shown in Figure 3A. Combining this $Y_B-R_m$ relationship and the equality in Equation (29), we show that the maximal growth rate occurs at a sub-optimal yield (Fig. 3B). Therefore, on the right-hand side of the maximal growth rate, there appears to be a trade-off between the growth rate and the yield. However, when measured under a wide range of growth conditions, the trade-off is not obvious. This result is consistent with *E. coli*’s choice of less-efficient fermentative pathway when glucose and oxygen are both abundant. This explanation does not exclude the argument of ATP yield optimization or the limitation of respiratory pathways. It does, however, provide another perspective.

The analysis shown here is based on Equations (1) and (2), which are generally accepted description for enzyme kinetics. Equation (1) states that substrate uptake rate is proportional to the copy number of the transporter multiplied by an unspecified kinetic function whose maximum value is unity. Therefore, the proportional constant is $k_{\text{cat}}$. Equation (2) is a balance between protein synthesis and protein degradation. These two equations are almost universally valid. Equation (12) is a direct mathematical consequence of Equations (1) and (2), and involves no additional assumption except that the cell is not undergoing a downshift in the transporter synthesis. With simple derivations, we uncovered a hidden boundary between growth rate and biomass yield, which escaped notice for many decades. This relationship is essential to the foundation of biochemical engineering, industrial biotechnology, as well as microbiology. Although more experiments are needed to test the generality and evolutionary implications, the simplicity of the analysis and its free of restrictive assumptions are quite appealing and promising.

![Figure 2. Square relationship between growth rate and biomass yield. A: The relative (denoted by subscript r) growth rate and the relative biomass yield in glucose followed the square-root relationship when cultured with or without amino acid (AA) supplementation or oxygen. The rates and yields were normalized to the wild-type values under aerobic conditions without AA supplements. All AA indicates all 20 amino acids were added to the culture while “some AA” indicates only some of the amino acids were added. B: The relative growth rate and the relative biomass yield followed the square-root relationship when cultured in various sugars with or without oxygen. The growth rate and biomass yield for each sugar is normalized against the aerobic condition in the same sugar.](image-url)
Figure 3. A: A typical effect of protein overexpression on yield. 
\( Y = 0.5 - (1 + \exp(-6Rm + 6))^{1/3} \). When protein synthesis rate \( Rm \) increases, the biomass yield decreases because of higher energy demand and potential saturation of downstream metabolic pathways. B: When the relationship in (A) is substituted in Equation (4), a maximum growth rate (open diamonds in A and B) occurs at a sub-optimal yield. Trade-off occurs when the growth rate and the yield of the strain starts on the right side of the maximum growth rate. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

References


