Fundamental *Escherichia coli* Biochemical Pathways for Biomass and Energy Production: Creation of Overall Flux States

Ross Carlson, Friedrich Srienc

Department of Chemical Engineering and Materials Science, and BioTechnology Institute, University of Minnesota, 240 Gortner Laboratory, 1479 Gortner Ave., St. Paul, Minnesota 55108; telephone: (612) 624-9776; fax: (612) 625-1700; e-mail: fried@cbs.umn.edu

Received 3 September 2003; accepted 4 December 2003

Published online 5 March 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/bit.20044

Abstract: We have previously shown that the metabolism for most efficient cell growth can be realized by a combination of two types of elementary modes. One mode produces biomass while the second mode generates only energy. The identity of the four most efficient biomass and energy pathway pairs changes, depending on the degree of oxygen limitation. The identification of such pathway pairs for different growth conditions offers a pathway-based explanation of maintenance energy generation. For a given growth rate, experimental aerobic glucose consumption rates can be used to estimate the contribution of each pathway type to the overall metabolic flux pattern. All metabolic fluxes are then completely determined by the stoichiometries of involved pathways defining all nutrient consumption and metabolite secretion rates. We present here equations that permit computation of network fluxes on the basis of unique pathways for the case of optimal, glucose-limited *Escherichia coli* growth under varying levels of oxygen stress. Predicted glucose and oxygen uptake rates and some metabolite secretion rates are in remarkable agreement with experimental observations supporting the validity of the presented approach. The entire most efficient, steady-state, metabolic rate structure is explicitly defined by the developed equations without need for additional computer simulations. The approach should be generally useful for analyzing and interpreting genomic data by predicting concise, pathway-based metabolic rate structures. © 2004 Wiley Periodicals, Inc.

Keywords: elementary mode analysis; metabolism; biochemical pathway; metabolic flux state; maintenance metabolism

INTRODUCTION

A number of methods have been developed to study the structure of biochemical networks and to probe typically unmeasurable, intracellular fluxes by analyzing the stoichiometric relationships of biochemical reactions (examples of recent reviews include Stephanopoulos et al., 1998; Lee and Papoutsakis, 1999; Schilling et al., 1999). For instance, with metabolic flux analysis one attempts to map an experimentally determined net rate pattern of consumed nutrients and of excreted metabolites to the individual reaction rates of a postulated intracellular reaction network. When experimental data are available for a range of operating conditions, one can gain some insight into the functioning of a cell from the changing intracellular reaction rate patterns at different growth conditions. On the other hand, if it is known which elementary modes (Schuster and Schuster, 1993; Schuster et al., 1994, 1999, 2000, 2002) a cell is using and at what rate it is growing, one can directly predict the net rates of nutrient consumption, by-product secretion, and all individual intracellular rates from the stoichiometry of the modes. The applied culturing conditions together with the operating elementary modes determine the mass transfer reactions between a cell and its environment. These transport processes are usually observable. Thus, one can use these observations to validate the operation of certain elementary modes in a cell.

We have previously demonstrated a method of efficiently sorting the dataset of all possible metabolic flux patterns identified by elementary mode analysis. This analysis identified unique, nondivisible pathways (elementary modes) operating in *E. coli* which most efficiently convert substrate into biomass and energy (Carlson and Srienc, 2004). We would like to refer to this previous study for a more comprehensive overview of previous work related to *E. coli* networks and elementary mode analysis; therefore, these topics are not repeated here. Experimentally observed intracellular regulation patterns as well as the metabolite secretion behavior in response to different degrees of oxygen limitation are in remarkable agreement with the operation of these most efficient elementary modes. In the current study, we extend this investigation by developing a method for computing the detailed, steady-state intracellular rate structure by treating the metabolism of a cell as a sum of two types of pathways: one pathway produces biomass and the second pathway operates solely to produce maintenance energy. This representation of the cellular metabolism, based on pathways identified in Carlson and Srienc (2004), permits the assembly and the explicit defining of a continuous,

Correspondence to: Friedrich Srienc

Contract grant sponsors: National Science Foundation, NIH training grant fellowship in Biotechnology (to R.C.)

steady-state, intracellular rate map based on a single set of computer simulation data. The predicted rate structure is compared with an extensive compilation of published, experimental data to support the validity of the approach and to demonstrate that *E. coli* cells likely grow using the most efficient pathways in spite of the many possible pathway choices. While the approach is validated by studying the well-documented glucose-limited *E. coli* system, the methodology is applicable to a wide variety of analyses, including network perturbations like the ones caused by gene knock-out mutations. Since elementary mode analysis identifies all possible steady-states pathways, the single set of data contains not only the optimal pathways for knock-out mutants but it also contains the pathways that represent the minimal host flux perturbation for knock-out mutants (Segre et al., 2002).

Pathway analysis approaches like elementary mode analysis simplify the evaluation of metabolic networks by organizing the repertoire of enzymatic reactions into meaningful, steady-state flux patterns. In contrast to methods that identify flux patterns through computations constrained by some optimization criteria, elementary modes are determined through the combinatorial analysis of all possible pathways. Elementary modes represent, therefore, the entire set of simplest, nondivisible flux units available to a network. The modes can provide insight into the capabilities and the behavior of biological organisms on the basis of fundamental metabolic pathways. Recently, blood cell metabolism has been examined in terms of extreme pathways which represent a very similar concept as elementary modes (Wiback and Palsson, 2002). This interesting analysis of a physiologically relevant network was limited in scope, however, because the relatively simple metabolic network of a red blood cell does not support growth. Furthermore, while elementary mode analysis and extreme pathway analysis are both based on convex analysis, limiting the set of identified pathways to the generating vectors of the convex cone can make the biological interpretation of extreme pathways more difficult than elementary modes. In particular, the red blood cell metabolic network considered reversible exchange reactions which likely limited the identified pathways to a smaller, less inclusive set than would be identified using elementary mode analysis (Schuster et al., 2002; Klamt and Stelling, 2003). The network presented in the current work includes one reversible exchange reaction (Carlson and Srienc, 2003). Therefore, the interpretation of the network using elementary mode analysis seems to be the more comprehensive approach. It extends our previous study of *E. coli* metabolism and presents a methodology based on elementary modes for analyzing and interpreting complex metabolic processes like cellular growth under conditions of culturing stress based on a minimal number of optimal pathways.

METABOLIC MODEL

The details of the metabolic model have been described previously (Carlson and Srienc, 2004). Briefly, the biochemical network model represents *E. coli* growth on glucose minimal media. The model includes 11 ''external'' metabolites, glucose, biomass, O_2, CO_2, acetate, formate, lactate, ethanol, succinate, NH_3, and a generic maintenance ATP term and includes 36 ''internal'' metabolites. The definition of ''internal'' and ''external'' metabolites can be found in Schuster et al. (2000). Of the 46 reactions included in the network, 18 are considered reversible, with the remaining 28 classified as irreversible. The model includes the central metabolism of *E. coli* and considers such features as glycolysis, the pentose phosphate pathway (PPP), the tricarboxylic acid (TCA) cycle, and the reactions involved in the formation of common metabolic by-products like acetate, ethanol, formate, lactate, and succinate. Growth rate-dependent biomass production was considered by creating different biomass terms to represent the seven studied growth rates. The configuration utilized in this study uses glucose as the sole energy source, while glucose and fixed CO_2 are both potential carbon sources.

The models were analyzed using the publicly available elementary mode analysis program METATOOL (v.352_double). This program is available at http://mudshark.brookes.ac.uk/sware.html or ftp://ftp.bioinf.mdc-berlin.de/Pub/metabolic/metatool/ (Schuster et al., 1994; Pfeiffer et al., 1999). The results were analyzed by pasting the output matrix into an MS Excel spreadsheet template. The spreadsheet simplified the sorting and plotting of results based on desired characteristics like carbon and oxygen yield or enzymes utilized. Each individual growth rate-dependent biomass term was run and analyzed as a separate model.

The model represents a steady-state chemostat culture, so the terms dilution rate and growth rate are used interchangeably. For the purposes of this article, the term elementary mode and biochemical pathway are considered interchangeable (Schuster et al., 2000).

THEORY

The accumulation rate of any cellular component, including substrates and products, can be described in the classical, chemical reaction network notation (Roels, 1983):

$$\mathbf{R} = \alpha\mathbf{M} \tag{1}$$

where $\mathbf{R}$ is the metabolite accumulation vector comprised of metabolite concentration time derivatives, α is the stoichiometry matrix of all reactions and metabolites in the network, and $\mathbf{M}$ is the velocity vector comprised of fluxes through each reaction (for references relevant to biological systems, see Stephanopoulos et al., 1998; Lee and Papoutsakis, 1999). The elements of the accumulation vector have a value of zero for metabolites that do not accumulate. The elements representing extracellular metabolites may or may not be zero, depending on the growth conditions. The typical objective of metabolic flux analysis (MFA) is to determine, based on measurable net rates $\mathbf{R}$ and the reaction stoichiometry represented by α, the individual reaction rates represented by the velocity vector $\mathbf{M}$.

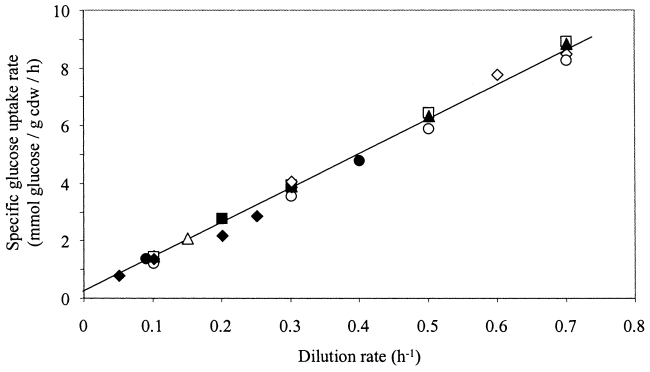


Figure 1. Specific glucose uptake rates (mmol/g cdw/h) vs. dilution rate (h^{-1}) for different strains of *E. coli* grown aerobically in glucose-limited chemostats. Linear regression of the data yielded the following equation relating dilution rate (h^{-1}) to specific glucose uptake rate (mmol/g cdw/h) of $r_{glc} = 12.14\ D + 0.17$. The data is taken from the following strains and references: unnamed strain (open circle) (Schultz and Lipe, 1964); strain ATCC 9001 (open square) (Neijssel et al., 1980); strain B/r and C(PC-1000) (filled triangle) (Tempest and Neijssel, 1987); strain B ATCC 11303 (filled square) (Snoep et al., 1993); strain ML30 (DSM 1329) (open diamond) (Lendenmann, 1994); strain MC4100 (open diamond) (de Graf et al., 1999); strain MG1655 (filled square) (Sauer et al., 1999); strain MC4100 (open triangle) (Alexeeva et al., 2000); strain W3110 (filled diamond) (Abdel-Hamid et al., 2001); strain JM101 (filled circle) (Emmerling et al., 2002). cdw = cell dry weight.

We have previously shown (Carlson and Srienc, 2003) that *E. coli* contains four unique pathways that most efficiently convert glucose into biomass and energy under any level of oxygenation. These pathways define four unique, discrete metabolic states MS1–MS4 which are used depending on the degree of oxygen limitation.

When cells grow in the most efficient manner, the vector **M** is composed of fluxes through the most efficient elementary modes. We have previously shown that the metabolism of a cell can be represented as the sum of two types of elementary modes when the cells grow in the most efficient manner. One mode operates for the synthesis of biomass and the second mode operates for the generation of cellular energy. Therefore, growth at the discrete metabolic states MS1–MS4 can be described as:

$$\mathbf{R}_i = \boldsymbol{\alpha}(\mathbf{M}_i^x + \mathbf{M}_i^{ATP}) \quad i = 1,2,3,4 \tag{2}$$

where $\mathbf{M}_i^x$ represents the fluxes from the biomass producing mode and $\mathbf{M}_i^{ATP}$ represents the fluxes from the energy producing mode. The subscript *i* refers to the discrete metabolic states. As shown previously, the energy mode consists of a subset of reactions that are used by the corresponding biomass mode. For each metabolic state, the flux through the pair of elementary modes specifies a defined accumulation vector for a given specific growth rate. One should note that this analysis represents just the inverse of the operation accomplished with conventional metabolic flux analysis, i.e., the observable system components found in vector **R** can be predicted based on the modes that are used.

Intermediate metabolic states can be obtained by taking linear combinations of the four discrete metabolic states. Such a strategy permits most efficient growth when there is an intermediate degree of oxygen limitation. The combination of modes can be formally expressed as:

$$\mathbf{R}_{i,i+1} = \boldsymbol{\alpha}[f_i(\mathbf{M}_i^x + \mathbf{M}_i^{ATP}) + (1 - f_i)(\mathbf{M}_{i+1}^x + \mathbf{M}_{i+1}^{ATP})]$$
$$i = 1,2,3 \tag{3}$$

where f_i and $(1-f_i)$ is the fractional use of the two metabolic states which serve as the endpoints of the line segment containing the intermediate metabolic state. A graphical representation of the line segment between the most efficient modes has been shown previously (Fig. 2, Carlson and Srienc, 2004). For instance, for any metabolic state between MS1 and MS2, $i = 1$; for any metabolic state between MS2 and MS3, $i = 2$; and for any metabolic state between MS3 and MS4, $i = 3$.

In practice, the elementary mode analysis was carried out with different stoichiometric coefficients for the biomass generating reaction in order to account for the growth rate-dependent changes in biomass composition (a detailed description of the growth rate dependent biomass composition can be found in Carlson and Srienc, 2004). Taking this into consideration, the stoichiometry matrix is a function of growth rate since the biomass-producing reaction stoichiometry changes with doubling time. The elementary modes obtained directly from the analysis are therefore somewhat arbitrary, although the ratios of the individual rates in the resulting modes are fixed. An easy conversion between the somewhat arbitrary modes and biologically significant rates like the specific growth rate are possible using a biomass carbon balance. While the macromolecular composition of *E. coli* is a function of growth rate, the biomass elemental carbon content is not. Most reported carbon contents for *E. coli* range from 45–50 mass percent. A sampling of published values is shown in Table I. These values do not vary with growth rate or substrate (Hempfling and Mainzer,

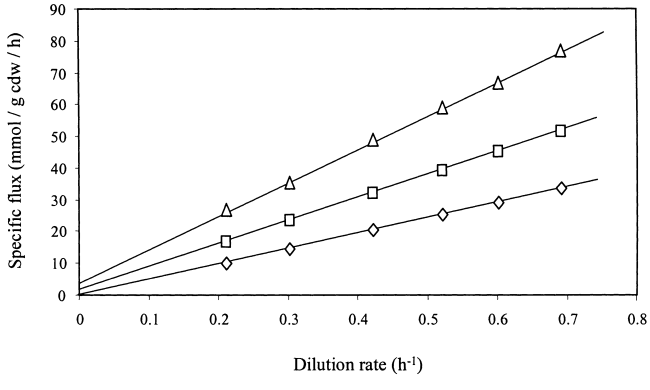


Figure 2. Specific carbon fluxes (Cmmol/g cdw/h) associated with glucose uptake (square) and biomass production (diamond) and the specific ATP production rate (mmol ATP/g cdw/h) (triangle) as functions of the dilution rate for oxygen-sufficient growth (MS1). The specific ATP production rate was calculated by assuming all glucose not used for biomass was used for energy production. The relationship between the specific ATP production rate and dilution rate is $r_{ATP} = 104.4\ D + 4.7$ mmol ATP/g cdw/h.

Table I. Carbon content of *E. coli* cell as percent of dry cell mass.

Strain	Carbon percentage	Reference
Unnamed	50 ± 5.0	Luria, 1960
B	49.1	Taylor, 1946
Unnamed	48.0 ± 0.2	Bratbak and Dundas, 1984
B	45.6 ± 2.5	Hempfling and Mainzer, 1975
MRE 600	45.5	Bauer and Ziv, 1976
B	45.4	Bauer and Ziv, 1976

For the presented study, it was assumed that carbon accounted for 48% of the cell dry weight.

1975). For purposes of this study, it was assumed that *E. coli* are 48% carbon by mass on a dry cell basis.

As noted above, the original elementary mode vector does not immediately give experimentally relevant rates, although the ratios between the individual rates are correct. More useful values are obtained by scaling the elementary mode vector with the following expression, which is based on a biomass carbon balance:

$$s_i^{x,\tau} = \left(\frac{\ln 2}{\tau}\cdot\frac{60\,\text{min}}{1\,\text{hour}}\right)\left(\frac{0.48\,\text{g carbon}}{1\,\text{g cdw}}\right)\left(\frac{1\,\text{Cmole}}{12\,\text{g carbon}}\right)\frac{1}{\alpha_{bio}^{x,\tau}}\frac{1}{m_{bio,i}^{\tau}}$$

$$i = 1,2,3,4 \tag{4}$$

This scaling factor is the ratio of the specific growth rate, expressed as Cmoles biomass/g cdw/h, divided by the number of Cmoles of biomass in the elementary mode. The scaling factor adjusts the relative velocity of the biomass mode into a biologically significant rate for a doubling time of τ (min). The metabolic state and doubling time are noted by the sub- and superscripts of s. The first set of brackets represents the specific growth rate (h^{-1}), the next two ratios convert the units to Cmoles biomass per g cdw per h: $\alpha_{bio}^{x,\tau}$ is the stoichiometric coefficient for biomass used in the elementary mode model, and $m_{bio,i}^{\tau}$ is the relative velocity of biomass synthesis in the biomass mode. $m_{bio,i}^{\tau}$ is an element of the elementary mode vector $\mathbf{M}_i^{x,\tau}$. The values for these terms, $\alpha_{bio}^{x,\tau}$ and $m_{bio,i}^{\tau}$, can be found in Appendix A.3 and Table II in Carlson and Srienc (2004), which describes the elementary mode computation in detail. When this scaling factor is multiplied by the biomass-producing elementary mode, an experimentally relevant accumulation vector related solely to biomass synthesis is obtained:

$$\mathbf{R}_i^{x,\tau} = \boldsymbol{\alpha}^{\tau}\mathbf{M}_i^{x,\tau}s_i^{x,\tau} \tag{5}$$

where $\mathbf{M}_i^{x,\tau}$ is the corresponding biomass elementary mode and $\boldsymbol{\alpha}^{\tau}$ is the doubling time-dependent stoichiometry matrix which accounts for the biomass composition. $\mathbf{R}_i^{x,\tau}$ contains the portions of the accumulation vector, $\mathbf{R}_i$, that are contributed by the biomass-producing elementary mode. For instance, it includes the specific glucose consumption rate, $r_{glc,1}^{x,\tau}$, which is devoted entirely to biomass synthesis. For this study, the dimensions of the term are moles glucose/g cdw/h. An example calculation, with units, is given in the Appendix.

The energy-producing elementary mode requires a different scaling factor. This factor can be determined by treating the total glucose consumption rate as a sum of two components. One component is a result of the biomass-producing elementary mode and the second component is a result of the energy-producing elementary mode:

$$r_{glc} = r_{glc}^{x} + r_{glc}^{ATP} \tag{6}$$

The specific glucose uptake rate under aerobic growth conditions has been experimentally determined in many investigations. These studies demonstrate that the specific glucose uptake rate, for each metabolic state, is a strong linear function of the specific growth rate. Therefore, one can use this experimental finding to estimate the scaling factor for the energy-producing elementary mode for metabolic state 1 (oxygen-sufficient conditions). It becomes:

$$s_1^{ATP,\tau} = (r_{glc,1}^{\tau} - r_{glc,1}^{x,\tau})\frac{1}{\alpha_{glc}}\frac{1}{m_{glc,1}^{ATP}} \tag{7}$$

where α_{glc} is the stoichiometric coefficient associated with glucose and $m_{glc,1}^{ATP}$ is the relative velocity of glucose consumption in the energy producing mode $\mathbf{M}_1^{ATP}$. Thus, the metabolic rate structure of a culture growing under conditions of metabolic state 1 for any growth rate is defined by:

$$\mathbf{R}_1 = \boldsymbol{\alpha}^{\tau}(\mathbf{M}_1^{x}s_1^{x,\tau} + \mathbf{M}_1^{ATP}s_1^{ATP,\tau}) \tag{8}$$

While this result is made possible by the extensively available glucose uptake rate data for aerobic growth, the extension of this result to other metabolic states requires further assumptions. We have assumed that the maintenance energy requirement for any metabolic state is the same and that it depends only on the specific growth rate. Therefore, a scaling factor can be generated that produces the same amount of ATP as the culture growing in metabolic state 1. To determine this scaling factor, it is convenient to define the ATP yield on glucose for an energy-producing elementary mode. This yield coefficient can be calculated from the ratio of appropriate elements found in the accumulation vector:

$$Y_{ATP/glu,i}^{ATP} = \frac{r_{ATP,i}^{ATP,\tau}}{r_{glu,i}^{ATP,\tau}} \quad i = 1,2,3,4 \tag{9}$$

This equation states that the ATP yield per glucose can be defined as the ratio of the specific ATP production rate to the specific glucose uptake rate. The scaling factor for the energy-producing mode can then be expressed in a more general way on the basis of the ATP yield of metabolic state 1 as:

$$s_i^{ATP,\tau} = \frac{Y_{ATP/glc,1}^{ATP}(r_{glc}^{\tau} - r_{glc,1}^{x,\tau})}{Y_{ATP/glc,i}^{ATP}}\frac{1}{\alpha_{glc}}\frac{1}{m_{glc,i}^{ATP}} \quad i = 1,2,3,4 \tag{10}$$

The energy mode scaling factor represents the ratio of the specific maintenance energy production rate to the relative velocity of ATP production in the energy mode. An example,

Table II. Scaling factors and predicted specific rates associated with each modeled growth rate under the four metabolic states.

Doubling time (min) metabolic state	$s_i^{x,\tau}$	$s_i^{ATP,\tau}$	r_{glc}	% C_x	r_{O2}	% O_{2x}	r_{ATP}	$r_{acetate}$	r_{EtOH}	$r_{formate*}$
200.1	1.22E-04	1.02	2.72	63	8.04	24	26.45	0.00	0.00	0.00
200.2	2.64E-04	2.20	4.23	48	5.87	25	26.45	5.61	0.00	0.00
200.3	3.96E-04	3.31	5.76	43	4.34	24	26.45	8.67	0.00	8.73
200.4	3.16E-03	8.82	13.00	32	0.00	0	26.45	11.56	11.58	23.20
100.1	4.16E-04	1.87	5.27	65	14.98	25	48.54	0.00	0.00	0.00
100.2	1.08E-02	4.04	8.07	50	10.98	26	48.54	10.39	0.00	0.00
100.3	2.70E-03	6.07	10.94	45	8.11	25	48.54	16.14	0.00	16.40
100.4	3.61E-03	16.18	24.46	34	0.00	0	48.54	21.55	21.63	43.44
80.1	2.45E-04	2.27	6.49	65	18.20	25	59.05	0.00	0.00	0.00
80.2	3.18E-03	4.92	9.87	50	13.35	26	59.05	12.56	0.00	0.00
80.3	1.59E-03	7.38	13.40	45	9.83	25	59.05	19.62	0.00	20.15
80.4	3.17E-03	19.68	29.76	34	0.00	0	59.05	26.16	26.20	52.89
60.1	2.73E-03	2.96	8.54	65	23.50	24	76.93	0.00	0.00	0.00
60.2	2.96E-03	6.41	12.87	50	17.29	26	76.93	16.12	0.00	0.00
60.3	4.44E-03	9.62	17.50	45	12.68	24	76.93	25.36	0.00	26.43
60.4	5.91E-03	25.64	38.60	34	0.00	0	76.93	33.79	33.79	68.65
50.1	6.33E-04	3.53	10.33	66	28.00	24	91.65	0.00	0.00	0.00
50.2	1.64E-02	7.64	15.47	51	20.65	26	91.65	19.07	0.00	0.00
50.3	4.11E-03	11.46	21.00	45	15.13	24	91.65	30.15	0.00	31.68
50.4	8.12E-03	30.55	46.18	34	0.00	0	91.65	40.21	40.31	82.05
40.1	1.07E-03	4.38	14.47	70	36.22	27	113.74	0.00	0.00	0.00
40.2	6.96E-03	9.48	21.04	55	26.81	29	113.74	24.43	0.00	0.00
40.3	3.48E-03	14.22	28.36	50	19.49	27	113.74	39.07	0.00	41.74
40.4	1.39E-02	37.91	60.87	38	0.00	0	113.74	52.07	51.99	106.74
30.1	3.03E-03	5.79	19.22	70	47.97	28	150.54	0.00	0.00	0.00
30.2	3.93E-02	12.55	27.93	55	35.52	29	150.54	32.33	0.00	0.00
30.3	9.84E-03	18.82	37.64	50	25.81	27	150.54	51.75	0.00	55.41
30.4	1.97E-02	50.18	80.66	38	0.00	0	150.54	68.95	68.84	141.45

The scaling can be used to determine the flux of any reaction in the metabolic network for growth rates with doubling times between 30 and 200 min and from oxygen-sufficient conditions to the complete absence of oxygen. The scaling factors and the listed specific rates all have units of mmol/g cdw/h. Note that none of these rates are in terms of Cmmol. See text for more details. glc = glucose, x = biomass, $\%C_x$ and $\%O_{2x}$ are the percentage of the glucose and oxygen fluxes, respectively, used for biomass production. Cdw = cell dry weight, EtOH = ethanol.

with units, is given in the Appendix. With the definition of the scaling factors, the metabolic rate structure of the growing cells at any point between metabolic state i and $i + 1$ can be described in a general way for the entire range of operating conditions with following expression:

$$\mathbf{R}_{i,i+1} = \boldsymbol{\alpha}^{\tau}\,[f_i(\mathbf{M}_i^{x,\tau} s_i^{x,\tau} + \mathbf{M}_i^{ATP} s_i^{ATP,\tau}) + (1-f_i)(\mathbf{M}_{i+1}^{x,\tau} s_{i+1}^{x,\tau} + \mathbf{M}_{i+1}^{ATP} s_{i+1}^{ATP,\tau})] \quad i = 1,2,3 \quad (11)$$

This equation determines all system fluxes by adding scaled combinations of a limited number of most efficient biochemical pathways. One set of pathways is used to produce biomass, while the second set of pathways is used to produce cellular energy for maintenance expenditures. The expression is analogous to Eq. (1), with the terms in the square brackets representing the vector of all individual network reaction fluxes which upon multiplication with the stoichiometry matrix results in the net accumulation rate vector **R**.

The following section illustrates how the scaling factors are determined from experimental aerobic glucose uptake rate data for a given specific growth rate. Specification of an additional nutrient uptake rate (such as the actual glucose uptake rate or oxygen uptake rate) or any metabolite secretion rate determines the two metabolic states involved and the weighting factor f and permits computation of the entire rate structure. A detailed sample calculation of a biomass and an energy mode scaling factor is given in the Appendix.

GLUCOSE UPTAKE RATES AND MAINTENANCE ENERGY

To estimate the pathway scaling factors, we analyzed glucose uptake rate data for aerobic and anaerobic growth conditions. The relationship between specific growth rate and glucose uptake rate for cultures not experiencing significant oxygen limitation has been measured in several studies

(for example, Schultz and Lipe, 1964; Hempfling and Mainzer, 1975). The results, summarized in Figure 1, show a strong linear relationship between the specific growth rate and the specific glucose uptake rate and show little variation between different *E. coli* strains.

These data can be used to estimate the magnitude of the flux through the energy-generating elementary modes. Since the growth rate-dependent carbon flux requirements for biomass production are known [see Eq. (5)], it is possible to determine the glucose flux associated with maintenance energy production using Eq. (6). This approach assumes that cells operate in the most efficient manner, as described by the MS1 fluxes.

The biomass glucose flux requirements were subtracted from the overall glucose uptake rate and the remaining glucose was assumed to be used for maintenance energy production. The specific ATP production rate was determined for each growth rate from the glucose uptake data and the MS1 energy mode stoichiometry [see Eq. (9)]. The maintenance energy requirements increase linearly with growth rate (see Fig. 2). This is not surprising, since the slope of the specific glucose uptake rate line is greater than the slope of the carbon flux directed toward biomass production. Extrapolating the specific ATP production rate to a dilution rate of zero suggests a 4.7 mmol ATP/g cdw/h nongrowth maintenance energy requirement which is similar to values reported previously (Schultz and Lipe, 1964; Hempfling and Mainzer, 1975; Tempest and Neijssel, 1987; Varma et al., 1993). The growth associated maintenance energy requirement is equal to 105.2·μ with units of mmol ATP/g cdw/h (Fig. 2). This expression is based on averaged experimental glucose uptake data (obtained from linear regression of data shown in Fig. 1) and the assumption that *E. coli* is 48% carbon by mass.

On the basis of the experimental glucose uptake data, the energy mode scaling factor applicable for metabolic state 1 can be calculated using Eq. (7). The net accumulation vector for MS1, $\mathbf{R}_1$, is then completely determined according to Eq. (8) for any given doubling time. We can then compare the predicted oxygen uptake rates for metabolic state 1 with experimental data to test the validity of the approach. Figure 3a shows that the predicted oxygen uptake rates fit well within the range of published experimental values. The close agreement between the predicted and the experimental values strongly supports the assumption that cells grow using the most efficient pathways.

The identified maintenance energy requirements are based on experimental data from cultures not experiencing significant oxygen limitation (MS1). We have assumed that the same maintenance energy requirement is also applicable for cultures experiencing oxygen stress. With this assumption, the scaling factors for the energy-producing modes can be calculated using Eq. (10). Since Eq. (11) is then completely specified for all conditions supporting a given specific growth rate, the net accumulation vector **R** is determined for any doubling time and any degree of oxygen limitation. On this basis, the anaerobic glucose uptake rates

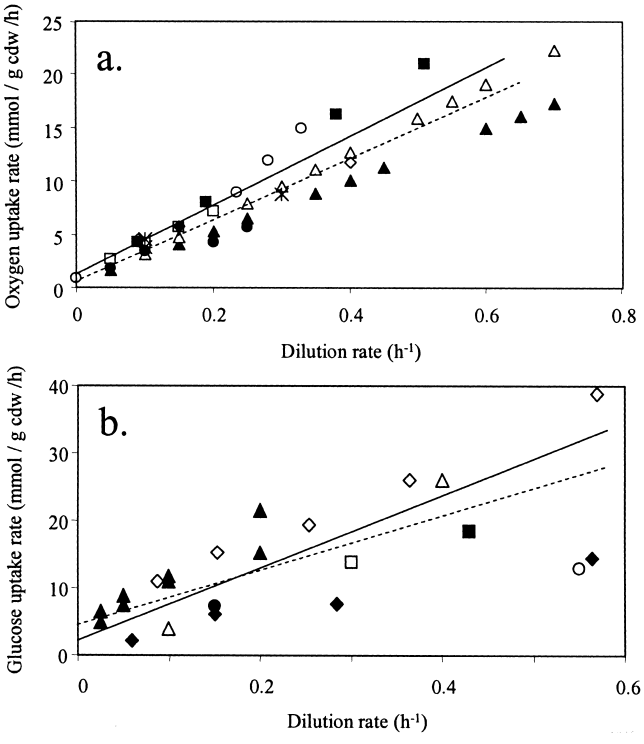


Figure 3. a: Predicted specific oxygen uptake rate (solid line) and linear regression fit of experimental oxygen uptake data (dashed line) vs. dilution rate. The published specific oxygen uptake rates from the following references: (filled triangle) (Schultz and Lipe, 1964); (filled square) (Hempfling and Mainzer, 1975); (open square) (Farmer and Jones, 1976); (open triangle) (Bajpai, 1987); (open diamond) (Calhoun et al., 1993); (filled diamond) (Snoep et al., 1993); (open circle) (Paalme et al., 1997); (star) (de Graf et al., 1999); (open diamond) (Sauer et al., 1999); (filled diamond) (Alexeeva et al., 2000); (filled circle) (Abdel-Hamid et al., 2001); (open diamond) (Emmerling et al., 2002). b: Predicted anaerobic specific glucose uptake rate (solid line) and linear regression fit of experimental anaerobic specific glucose uptake rate (dashed line) both as mmol glucose/g cdw/h vs. dilution rate. The plotted data came from the following references: (open diamond) (Hempfling and Mainzer, 1975); (filled diamond) (Chesbro et al., 1979); (open square) (Snoep et al., 1993); (filled square) (Varma and Palsson, 1994); (filled circle) (Alexeeva et al., 2000); (open triangle) (Berrios-Rivera et al., 2000); (open circle) (Emmerling et al., 2002); (filled triangle) (Riondet et al., 2000). cdw = cell dry weight.

were predicted as a function of the growth rate for an anaerobic culture (MS4) [Eq. (11); i = 3; $f_3 = 0$]. These data are plotted together with published experimental data in Figure 3b. The prediction is well within a single standard deviation of the experimental values. It therefore appears reasonable to assume that the maintenance energy requirements are the same regardless of oxygen stress and that cells likely operate in the most efficient way when they grow anaerobically. These findings support results obtained with linear programming methods (Edwards et al., 2001).

DEFINING THE METABOLIC RATE STRUCTURE

Scaling factors that convert the most efficient modes into physiologically appropriate rates were developed in the pre-

vious section. These scaling factors depend only on the specific growth rate. One can then compute, for a given specific growth rate, all four possible optimal metabolic states between completely aerobic and completely anaerobic growth conditions. Table II lists the scaling factors that define the metabolic rate structure based on elementary modes for most efficient steady-state glucose-limited *E. coli* growth under varying levels of oxygen stress. The table, together with Eq. (11), can be used to explicitly define every point in the steady-state solution space without further computer simulations. Variation of the fractional contribution, *f*, permits computation of growth conditions between any two adjacent metabolic states. Alternatively, for a known growth rate, specification of a glucose or oxygen uptake rate (or any by-product transfer rate) fixes the identity of the two contributing metabolic states and their fractional contributions. The optimal behavior of a glucose-limited cell culture can be expressed by the simple nonnegative, linear combination of only four unique metabolic states.

It is possible to plot the relationship between any cellular fluxes without further computer simulations using the defined rate structure. For instance, Figure 4 shows a three-dimensional relationship between glucose fluxes, oxygen fluxes, and growth rate. This figure is analogous to a phenotypic phase plane (Edwards et al., 2002); however, the basis of the presented work is the stoichiometry from nondivisible pathways (Carlson and Srienc, 2004) and not the mathematical concept of a ''shadow price'' used in linear programming. There are only two degrees of freedom for the steady-state solutions represented in Figure 4. Fixing two rates sets the value of the third parameter. The fluxes show that metabolic state 2 occurs when the available oxygen is about 74% of the level required for oxygen-sufficient growth (MS1) at a given doubling time. The specific oxygen uptake rates for each metabolic state and for each considered growth rate are also listed in Table II. For instance, the specific oxygen uptake rate for a culture with an 80-min doubling time is 18.2 and 13.4 mmol O_2/g cdw/h for MS1 and MS2, respectively. Knowing the doubling time and oxygen uptake rate sets the value of the third parameter: specific glucose uptake rate. A specific oxygen uptake rate of 13.4 mmol O_2/g cdw/h and a doubling time of 80 min fix the specific glucose uptake rate at 9.9 mmol glucose/g cdw/h. Metabolic state 3 occurs when the available oxygen is about 54% of the level required for oxygen-sufficient growth at a given doubling time. The increased requirement for glucose under conditions of oxygen stress is also evident in Figure 4. Completely anaerobic growth conditions require ~4.5 times more glucose than oxygen-sufficient conditions. The slopes of the lines representing a constant doubling time indicate the sensitivity of the *E. coli* network to oxygen and offer an approach for determining metabolic response coefficients with respect to oxygen. The region between metabolic states 3 and 4 is especially sensitive to small changes in the oxygen fluxes. In this region, small amounts of oxygen have a large effect on the required glucose flux. Since each line on the doubling time axis represents the same amount of biomass,

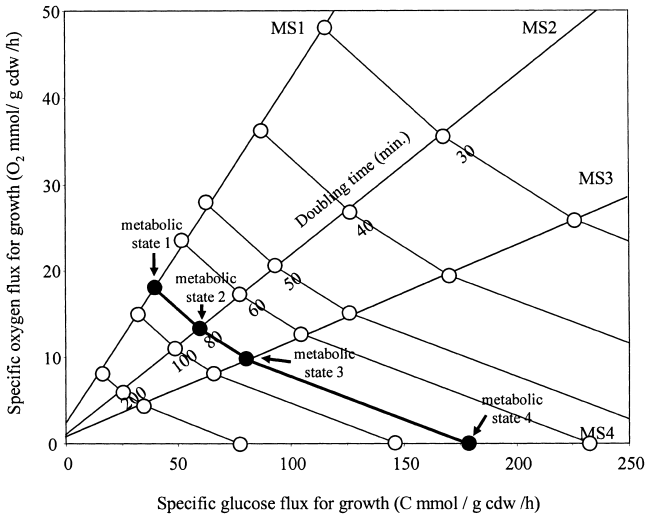


Figure 4. The 3D relationship between specific oxygen flux, specific glucose carbon flux, and specific growth rate presented as a doubling time for most efficient, glucose-limited growth of *E. coli*. The circles represent the most efficient modes for spanning the growth conditions from anaerobiosis to completely aerobic growth. The doubling time axis is labeled on top of the line representing metabolic state 2 (MS2). The isoline representing a culture with a doubling time of 80 min is highlighted with a heavy line. Every point on this line represents a culture with an 80-min doubling time under different levels of oxygen stress. The metabolic state lines are linear regression fits of the data points shown in the plot as open circles. The line fit r^2 values were better than or equal to 0.997 for all three lines. This chart considers both biomass production and maintenance energy requirements. The metabolic state lines do not go through the origin because of the nongrowth associated maintenance energy requirements. Any cellular flux associated with optimal growth can be determined using Table II and Appendixes A.4 and A.2 in Carlson and Srienc (2004). All other locations in the operating space can be defined by linear combinations of the nearest data points. The same type of chart can be constructed to show the relationship between any two intracellular fluxes and the growth rate.

an increase in glucose uptake rate corresponds linearly with a decrease in biomass carbon yield. This information could be important for such calculations as determining the relative cost benefits of limited aeration on media substrate costs.

The rate structure contains a large amount of information related to the most efficient conversion of nutrients into biomass and maintenance energy. The following examples illustrate some principles and applications of this solution space to answer physiological questions. The predictions are compared with experimental results, when available, as an additional means of validating the modular, pathway-based interpretation of the *E. coli* metabolism.

BY-PRODUCT SECRETION PROFILES

The most efficient rate structure was used to find the by-product secretion profile of an *E. coli* culture growing at a constant growth rate with varying degrees of oxygen limitation. The model predictions are compared in Figure 5 with a recent experimental study by Alexeeva et al. (2000) which examined the same question. The pathway approach pro-

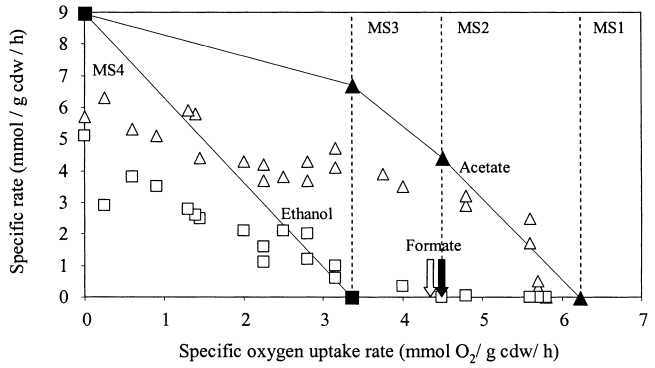


Figure 5. Predicted specific acetate production rates based on ''typical'' *E. coli* (filled triangle) and published specific acetate production rates from Alexeeva et al. (2000) (open triangles). Predicted ethanol-specific production rates based on ''typical'' *E. coli* (filled squares), and published ethanol production rates from Alexeeva et al. (2000) (open squares). The arrows show the conditions of predicted formate product (filled arrow) and the experimental point of formate production (open arrow).

vides a rational, relatively simple basis for interpreting the metabolic switches in *E. coli* metabolism at various levels of oxygen stress.

The analysis predicts very well the switching between the four metabolic states as a function of oxygen limitation as shown by the onset of excretion of the byproducts acetate, formate, and ethanol. While the qualitative trend of secretion rates is also predicted, the computed values deviate somewhat from the experimental values when growth becomes increasingly anaerobic. There are several possibilities that can explain this deviation. First, experimental data may not be accurate, as it is difficult to avoid evaporation of the volatile compounds. This could significantly reduce experimentally measured values. Second, the predictions are based on the same maintenance energy demand under anaerobic conditions as evaluated for aerobic conditions. A lower maintenance energy requirement under anaerobic conditions would decrease by-product secretion rates. However, this is in contrast to the glucose consumption rate data shown in Figure 3b, which supports the assumption of the same maintenance energy requirement. Third, the cells may not use the predicted most efficient pathways under anaerobic conditions. Under some conditions, processes like the diffusion of acetate or ethanol out of the cell may limit the flux through the most efficient mode. Under such conditions, a slightly less efficient mode which uses other electron sinks, like succinate, could be used in addition to the most efficient mode. As previously shown (Carlson and Srienc, 2003), the *E. coli* network contains a number of elementary modes that are only slightly less efficient under anaerobic conditions. Some of these modes secrete succinate. Secretion of succinate could decrease acetate and ethanol secretion rates. The culture studied by Alexeeva et al. (2000) produced some succinate but only trace amounts of lactate. In addition to being consistent with experimental data, these results are also consistent with the by-product secretion order predicted in Varma et al. (1993) using LP methods, further validating the presented approach.

OXYGEN CONSUMPTION AND LIMITATION IN RAPIDLY GROWING CULTURES

Rapidly growing cultures are important for biotechnology processes, but they often experience oxygen limitation. The onset of partially oxidized by-product secretion is believed to be the result of insufficient oxygen transfer to the cells. The oxygen transfer process is a complex function of oxygen solubility, oxygen diffusion rates, and cell geometry (i.e., the surface area to volume ratio) (Andersen and von Meyenburg, 1980). The reported maximum oxygen uptake rates for a number of *E. coli* strains growing in glucose minimal media can be found in Table III. Most reported values are between 15 and 21 mmol O_2/g cdw/h. Depending on the strain and the culturing conditions, the onset of partially oxidized by-product production can differ. As the metabolic demand for oxygen surpasses the maximum uptake rate, the cell switches from an oxygen-sufficient metabolism (MS1) toward an oxygen-limited metabolism (MS2 and MS3). The transition from MS1 to MS2 results in the production of acetate. The growth rate at which oxygen limitation first occurs has significant implications on the behavior of a culture. The transition from oxygen sufficiency to oxygen limitation requires an increase in glucose catabolism (see Fig. 4). The predicted specific glucose uptake rates for cells with a maximum specific oxygen uptake rate of 15 mmol/g cdw/h and a maximum rate of 20 mmol/g cdw/h are shown in Figure 6.

Figure 6a shows both the glucose and oxygen uptake rates vs. growth rate for a culture with a maximum oxygen uptake rate of 15 mmol O_2/g cdw/h. The culture can grow using MS1 until a growth rate of $\sim 0.43\ h^{-1}$. At this point, oxygen demand surpasses the maximum possible uptake rate and the cultures start growing with an intermediate metabolic state between MS1 and MS2. Oxygen is continuously available; however, it is only supplied at a rate of 15 mmol O_2/g cdw/h. The transition from oxygen-sufficient to oxygen-limited growth lowers both the biomass and ATP yield on glucose, so the glucose consumption rate increases at a faster rate with respect to the specific growth rate, as seen by the change of slope in Figure 6a. The transition to an oxygen-limited metabolic state results in the production of acetate. At a growth

Table III. Maximum experimental oxygen uptake rates for *E. coli* grown in glucose minimal media.

E. coli strain	Maximum O_2 uptake rate (mmol O_2/g cdw/h)	Reference
B	21	Hempfling and Mainzer, 1975
B/r	20	Andersen and von Meyenburg, 1980
RB791	20	Lin et al., 2000
W3110	19	Lin et al., 2000
Unnamed	16	Schulze and Lipe, 1964
W3110	15	Varma and Palsson, 1994

cdw = cell dry weight.

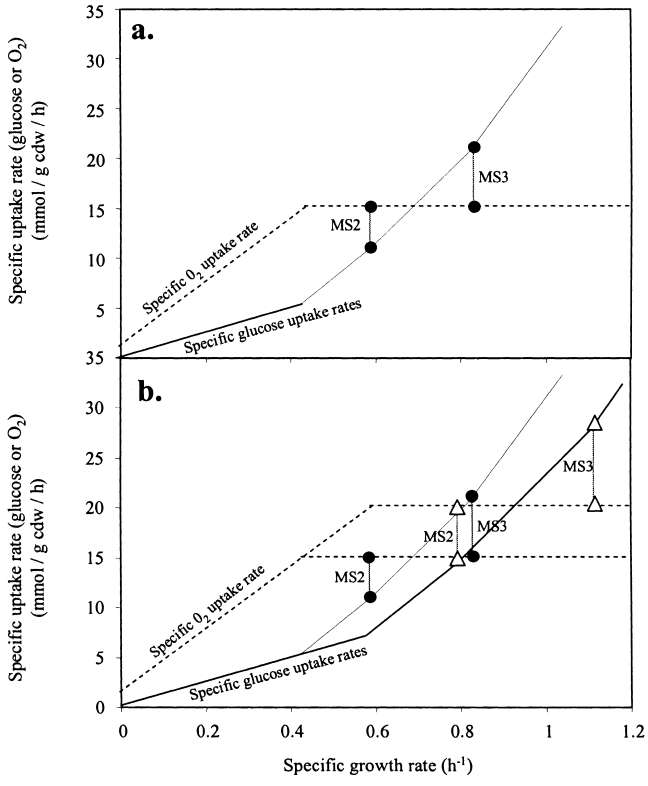


Figure 6. a: Predicted specific glucose (solid lines) and specific oxygen uptake rates (dotted lines) vs. growth rate. The plot shows the predicted glucose uptake rates for cultures having a maximum oxygen uptake rate of 15 mmol O_2/g cdw/h. Under conditions of oxygen limitation, the glucose uptake rate increases to maximize growth efficiency for the available oxygen. The points labeled MS2 and MS3 correspond to culturing conditions defined as metabolic states 2 or 3, respectively. See text for more details. The maximum specific oxygen uptake rate also indicates when the cultures start to produce significant levels of the partially oxidized by-product acetate, formate, or ethanol. For instance, with a maximum oxygen uptake rate of 15 mmol/g cdw/h, acetate would be produced starting at a dilution rate of $\sim 0.43\ h^{-1}$. b: Comparison of glucose uptake rates vs. growth rate for cultures with a maximum oxygen uptake rate of either 15 (circles) or 20 (triangles) mmol O_2/g cdw/h. See text for more details. cdw = cell dry weight.

rate of $\sim 0.58\ h^{-1}$, the culture reach MS2. Another change in the slope of the glucose uptake rate is seen as the cells transition toward intermediary metabolic states between MS2 and MS3. When the culture surpasses the conditions defined by MS2, the cells begin to produce formate. As mentioned earlier, MS2 occurs when the supplied oxygen is $\sim 74\%$ of the total oxygen required for oxygen-sufficient growth and MS3 occurs when the supplied oxygen is about 54% of the oxygen needed for oxygen-sufficient growth.

The maximum glucose uptake rate and the maximum oxygen uptake rate play an important role in culture behavior. The smaller the maximum oxygen uptake rate, the more glucose a culture will need to maintain a steady state at a given growth rate. For instance, a culture with a maximum oxygen uptake rate of 15 mmol O_2/g cdw/h needs $\sim 40\%$ more glucose than a culture with a maximum O_2 uptake rate of 20 mmol O_2/g cdw/h to maintain a steady-state culture at a growth rate of $0.58\ h^{-1}$ (see Fig. 6b). This corresponds to a difference in biomass and ATP substrate yields of $\sim 40\%$. The maximum oxygen uptake rates likely play a role in a culture's maximum growth rate. The maximum reported glucose uptake rates for an aerobic culture are around 10.5 mmol glucose/g cdw/h (Varma and Palsson, 1994; Lin et al., 2001); therefore, the maximum specific growth rate for cultures with the considered maximum oxygen uptake rates would be around $0.58\ h^{-1}$ to $0.74\ h^{-1}$. These growth rates are in the range reported by Senn et al. (1994) in a survey of maximum specific growth rates for *E. coli* growing on glucose minimal media in both batch and chemostat cultures.

The effect of oxygen limitation on culture parameters like specific acetate production rates can be directly predicted using the information in Table II. For instance, a culture with a maximum oxygen uptake rate of 15 mmol/g cdw/h reaches MS2 at a dilution rate of $0.58\ h^{-1}$, which is equivalent to a doubling time of ~ 72 min (see Fig. 6). The specific acetate production rate at this point can be determined by using a linear average of the two closest growth rates found in Table II. The specific acetate production rate ($r^{\tau=80}_{acetate,2}$) for a culture in MS2 and having an 80-min doubling time ($\mu = 0.52\ h^{-1}$) is 12.6 mmol/g cdw/h, while a culture with a 60-min doubling time ($\mu = 0.69\ h^{-1}$) in MS2 has a specific acetate production rate ($r^{\tau=60}_{acetate,2}$) of 16.1 mmol/g cdw/h. The contribution from each growth rate is determined using the lever rule and can be described with the following expression:

$$r^{\tau=72}_{acetate} = r^{\tau=80}_{acetate}\left(1 - \frac{d_1}{d_1 + d_2}\right) r^{\tau=60}_{acetate,2}\left(1 - \frac{d_1}{d_1 + d_2}\right) \quad (12)$$

where d_1 and d_2 are the line lengths taken from point ''A'' in Figure 7. The predicted specific acetate production rate is 14.0 mmol/g cdw/h for a culture with a maximum O_2 uptake rate of 15 mmol/g cdw/h and a growth rate of $0.58\ h^{-1}$. The lever rule can also be used to determine the contribution from the two closest metabolic states. For instance, with a

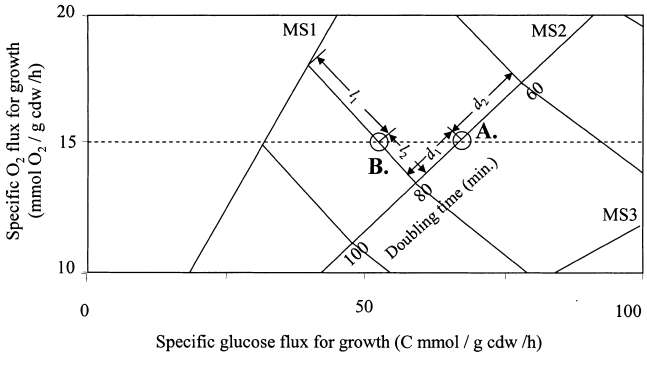
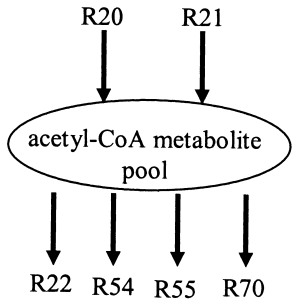


Figure 7. Illustration of the lever rule for determining flux patterns for cultures with doubling times not specifically studied (point A) and for intermediate metabolic states (point B). By using linear averages of the nearest defined points, it is possible to determine the intracellular fluxes for any position in the continuous operating space. See text for more details. The plot is a subsection from Figure 4. cdw = cell dry weight.



net specific enzymatic flux (mmol /g cdw / h)

	R20	R21	R22	R54	R55	R70
MS1.	0	3.10	2.87	0	0	0.22
MS2.	0	6.10	0.27	0	5.60	0.22
MS3.	9.18	0	0.27	0	8.68	0.22
MS4.	23.64	0	0.27	11.58	11.56	0.22

Figure 8. Metabolite flux partitioning around acetyl-CoA for a steady-state *E. coli* culture with a 200-min doubling time under different levels of oxygenation. Under different levels of oxygen stress, the flux rates into and out of the acetyl-CoA metabolite pool change significantly. The designations of the reactions can be found in Appendix A.1 of Carlson and Srienc (2004); see text for more details. R20 = pyruvate-formate lyase; R21 pyruvate decarboxylase complex; R22 = citrate synthase; R54 = acetaldehyde dehydrogenase/alcohol dehydrogenase, R55 = phosphotransacetylase/ acetate kinase and R70 = biomass synthesis requirements. cdw = cell dry weight.

maximum oxygen uptake rate of 15 mmol O_2/g cdw/h and a doubling time of 80 min (point ''B'' in Fig. 7), the predicted specific acetate production rate is defined by:

$$r_{acetate} = r^{\tau=80}_{acetate,1}\left(1 - \frac{l_1}{l_1 + l_2}\right) + r^{\tau=80}_{acetate,2}\left(1 - \frac{l_2}{l_1 + l_2}\right) = r^{\tau=80}_{acetate,1} f_1 + r^{\tau=80}_{acetate,2}(1 - f_1) \qquad (13)$$

where l_1 and l_2 are the line lengths for point ''B'' in Figure 7 and the terms $r^{\tau=80}_{acetate,1}$ and $r^{\tau=80}_{acetate,2}$ are the specific acetate production rates for a culture with a doubling time of 80 min in MS1 and MS2, respectively. The expression also demonstrates the use of the fractional contribution factor f_1. This factor sets the nonnegative, fractional contribution of the two nearest metabolic states, which in this case are MS1 and MS2. Note that there is no contribution from MS3 or MS4. The predicted specific acetate production rate under these conditions is 8.4 mmol acetate/g cdw/h.

METABOLITE FLUXES AND PARTITIONING AT BRANCH POINTS

The defined rate structure contains the flux through every examined reaction and therefore permits analysis of metabolite flux partitioning at all branch points under different doubling times and under different levels of oxygen stress. The partitioning of metabolite fluxes around the acetyl-CoA pool for a culture growing with a 200-min doubling time is shown in Figure 8. The metabolic model accounts for two different acetyl-CoA synthesis reactions: pyruvate dehydrogenase complex (R21) and pyruvate formate lyase (R20); and four reactions that consume acetyl-CoA: citrate synthase (R22), acetaldehyde dehydrogenase/alcohol dehydrogenase (R54), phosphotransacetylase/acetate kinase (R55), and biomass synthesis requirements (R70) (for stoichiometries, see Appendix A.1 in Carlson and Srienc, 2004).

To operate most efficiently under different levels of oxygen stress, the enzymes shown in Figure 8 must have different activities under different conditions. For instance, the flux through citrate synthase (R22) decreases more than 10-fold between MS1 and MS2, the flux through phosphotransacetylase/acetate kinase (R55) increases more than 2-fold between MS2 and MS4, and the acetyl-CoA drain for biomass production (R70) remains the same, regardless of oxygen stress. The analysis gives insight into the structure of the regulatory network by predicting what metabolite fluxes are most efficient for a given set of culturing conditions. The metabolite flux information in Figure 8 could be systematically analyzed using the tools of metabolic control analysis where the relationship between fluxes and enzyme activities are analyzed (for a review see Kacser et al., 1995). The information could also be useful for creating a structured basis to design and to interpret the results of ^{13}C-based flux studies (for recent review, see Wiechert, 2001). This type of analysis can be done at any branch point in the central metabolism.

DESCRIPTION OF DISTRIBUTED CELL POPULATIONS

Cell cultures are not typically comprised of a homogeneous population. There is often significant heterogeneity of parameters like cell size, microenvironments, age, or genetic make-up. Differences in parameters like cell size could potentially lead to a distribution of other parameters like the maximum specific oxygen uptake rates due to different surface area to volume ratios (Andersen and von Meyenburg, 1980). The presented analysis provides a structured basis for interpreting potential ramifications of population heterogeneity on culture parameters. Techniques like flow cytometry (FCM) are capable of measuring single cell properties. A recent study has shown that *E. coli* populations have a distribution of glucose uptake rates (Natarajan and Srienc, 1999, 2000). A hypothetical distribution of glucose uptake rates and its relationship to growth rate and oxygen uptake rate is shown in Figure 9. It is clear from this figure that a cell population exhibiting heterogeneous glucose uptake rates

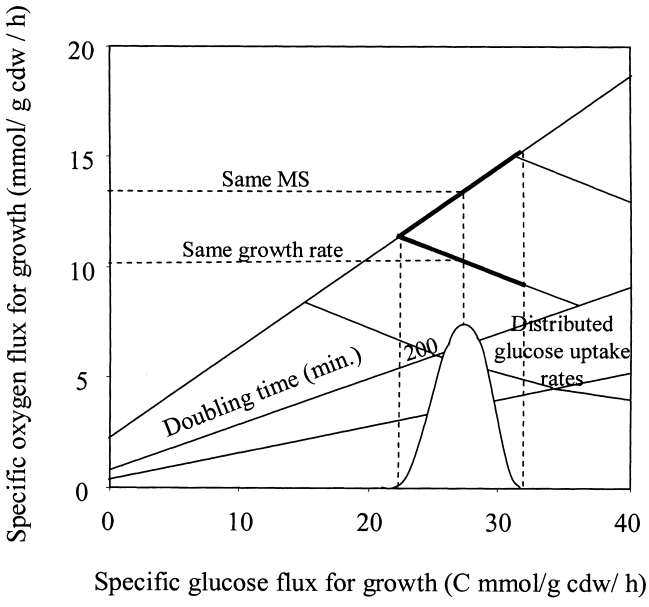


Figure 9. Two potential scenarios based on a population distributed specific glucose uptake rate. A hypothetical distribution of glucose uptake rates, shown along the x-axis, are projected onto the MS1 line suggesting a distribution of growth rates and projected onto a single growth rate line suggesting a distribution of metabolic states. The plot is a subsection of Figure 4. cdw = cell dry weight.

must be heterogeneous in other properties as well. Such a cell population likely has a distribution of either growth rates and/or of metabolic states. But the real situation may be even more complicated if cell-to-cell interactions can take place in which metabolites are secreted by some cells and consumed by others.

DISCUSSION

In the presented study, we develop a method of assembling steady-state solution spaces from a minimal number of unique, nondecomposable pathways. This is done by treating the intracellular fluxes as a combination of separate fluxes for biomass synthesis and maintenance energy production. By defining the doubling time-dependent maintenance energy requirements, it is possible to assemble the complete rate structure by adding weighted, nonnegative linear combinations of a limited number of these pathways. The validity of the results was tested by comparing the predictions with the extensively available experimental and theoretical data related to glucose-limited growth under varying degrees of oxygen stress. The approach was used to provide a mechanistic basis for interpreting experimental results like the relationship between maximum specific oxygen uptake rates and the maximum specific growth rates.

The relationship between elementary mode analysis, extreme pathway analysis, and linear optimization techniques has been described in detail in several articles (Schilling et al., 1999; Schuster et al., 2002; Klamt and Stelling, 2003). All three techniques operate on the same mathematical basis: a convex cone that contains the steady-state solution space. Briefly, elementary mode analysis identifies all possible, unique, nondivisible pathways available to a network, including the generating pathways for the convex cone. Extreme pathway analysis, on the other hand, identifies only the convex cone generating pathways. Limiting the analysis to just the generating vectors can complicate analysis by excluding biologically significant or important pathways. For instance, the most efficient biomass elementary mode for completely aerobic growth (MS1) is not a generating pathway. Traditional linear programming (LP) methods typically find a point on the surface of the convex cone that satisfies a set of subjective optimality criteria (Savinell and Palsson, 1992; Varma and Palsson, 1993). Depending on the applied optimization criteria, these traditional LP methods do not always identify unique, nondivisible pathways. Instead, they may identify flux states that are comprised of linear combinations of elementary modes. Recently, a mixed integer linear programming (MILP) technique has been developed which can identify all ''alternate optima'' for a user-supplied optimality criteria (Lee et al., 2000; Phalakornkule et al., 2001). It would be interesting to find out to what extent these optima coincide with elementary modes, since this approach could potentially represent an alternate way to compute fundamental pathways. While elementary mode analysis can require significant computational effort to determine all pathways for a complex metabolic network, the output file represents the complete set of physiologically meaningful pathway possibilities (Schuster et al., 2002), which makes the application of this concept very attractive.

The anaerobic glucose uptake data in Figure 3b shows significantly more scatter than the aerobic data shown in Figure 1. This can potentially be explained by examining the by-products of anaerobic growth. The most efficient means of anaerobic *E. coli* growth, as predicted by MS4, results in the formation of acetate, ethanol, and formate, but several reports indicate that anaerobic *E. coli* cultures often produce varying amounts of lactate and succinate (Blackwood et al., 1956; Belaich and Belaich, 1976; Chesbro et al., 1979; Alam and Clark, 1989; Alexeeva et al., 2000). Mode analysis suggests that these by-products are the result of slightly less efficient anaerobic energy-producing pathways (data not shown). Under some conditions, the pathways associated with M_4^{ATP} may limit the rate of ATP production, requiring the cell to utilize other, less efficient lactate and succinate-producing energy pathways to meet the maintenance energy requirements. To make up for the lower ATP yields of these pathways, an increased rate of glucose catabolism would be required. The flux through the less efficient modes is likely a strain and a condition-dependent effect. Different articles report very different levels of these two by-products (Blackwood et al., 1956; Belaich and Belaich, 1976; Chesbro et al., 1979; Alam and Clark, 1989; Alexeeva et al., 2000). Some of the experimental data points in Figure 3b are also slightly below the predicted glucose uptake rate. The glucose uptake rates in near anaerobic *E. coli* cultures are quite sensitive to the presence of oxygen or other electron

acceptors. Even small amounts of contaminating oxygen can have a large effect on the glucose uptake rates (see Fig. 4). Nonideal experimental conditions that have residual oxygen or other electron acceptors could potentially lower the observed glucose uptake rates.

The metabolic network structure defined in the presented study is based on glucose minimal media and examined oxygen stress. Other carbon sources like pentose sugars and different culturing stresses like nitrogen or phosphorous limitation would have different parameters like ATP or biomass carbon yields. Defining the operating spaces for these different conditions could be very useful for analyzing and interpreting results from DNA microarray, proteomic, or metabolic flux experiments. The methodology established in the presented study demonstrates how metabolic solution spaces can be constructed from physiologically meaningful networks to study processes as complex as cellular growth under varying levels of culturing stress. Similar models for other organisms like *Saccharomyces cerevisiae* could also be created from the available biochemical data and annotated genomic sequences.

APPENDIX

The following is an example for determining the biomass and energy scaling factor for an aerobic culture (MS1) with a 200-min doubling time. The external metabolite stoichiometry for the biomass and energy mode are shown below. These pathways are taken from Carlson and Srienc (2004):

$$14,005\ \text{glucose} + 19,006\ NH_3 + 15,806\ O_2 = 15,078\ CO_2 + 26\,\text{biomass} \tag{A.1}$$

$$\text{glucose} + 6\ O_2 = 26\,\text{ATP} + 6\ CO_2 \tag{A.2}$$

The scaling factor for the biomass pathway can be determined using Eq. (4). The term $\alpha_{bio}^{x,200}$ is the stoichiometric coefficient that converts the biomass term into a Cmole basis. In this case, it is equal to 2652, which can be found in Table II of Carlson and Srienc (2004). The term $m_{bio,1}^{200}$ is the relative biomass production rate and is equal to 26 Eq. (A.1). Therefore, the scaling factor is:

$$s_1^{x,200} = \left(\frac{\ln 2}{200\,\text{min}} \cdot \frac{60\,\text{min}}{1\text{hr}}\right)\left(\frac{0.48\,\text{g carbon}}{1\,\text{g cdw}}\right)\left(\frac{1\,\text{Cmole}}{12\,\text{g carbon}}\right)\frac{1}{2652\ \text{Cmole}}\frac{1}{26} = \frac{1.22\,\text{x}10^{-7}}{\text{g cdw hr}} \tag{A.3}$$

As mentioned previously, this assumes biomass is 48% carbon by mass. The weighting factor in Eq. (A.3) is simply the ratio of the specific growth rate in terms of Cmoles biomass/g cdw/h divided by the number of Cmoles in the biomass mode found in Eq. (A.1). Multiplying the scaling factor by any element of the biomass synthesizing mode vector gives a biologically significant rate. For instance, multiplying $s_1^{x,200}$ by the glucose term [14,005 moles in Eq. (A.1)] gives a specific glucose consumption rate for biomass synthesis ($r_{glc,1}^{x,200}$) of 1.71×10^{-3} mol glucose/g cdw/h. The biologically significant rates are not limited to the external metabolites. The rate through any intracellular reaction is also biologically significant.

The scaling factor for the energy mode can be determined using Eq. (10). The ATP yield per glucose is 26 based on the stoichiometry of Eq. (A.3). Figure 1 gives a specific glucose uptake rate ($r_{glc,1}^{200}$) of 2.72×10^{-3} mol/g cdw/h for an aerobic culture with a 200-min doubling time. It was shown above that the biomass synthesis requires a specific glucose uptake rate of 1.71×10^{-3} mol/g cdw/h. The term α_{glc} is the stoichiometric coefficient for glucose, which is 1 for all energy modes. The $m_{glc,1}^{ATP}$ term is the relative rate of glucose catabolism for the MS1 energy mode, which is 1 based on Eq. (A.2). Therefore, the scaling factor is:

$$s_1^{ATP,200} = \left(\frac{26\,\frac{ATP}{glc} \cdot \frac{1.01\times10^{-3}\text{mole glc}}{\text{g cdw hr}}}{26\,\frac{ATP}{glc}}\right)\frac{1}{1\,\text{mole glc}} \cdot \frac{1}{1} = \frac{1.01 \times 10^{-3}}{\text{g cdw hr}} \tag{A.4}$$

Equation (A.4) simply represents the doubling time-dependent specific ATP production rate found in Figure 2 divided by the number of ATP produced in the MS1 energy mode. This weighting factor converts every term in Eq. (A.2) into a biologically significant rate. For instance, multiplying $s_1^{ATP,200}$ by the ATP term suggests a culture growing aerobically with a 200-min doubling time has a specific maintenance energy production rate of 26.3×10^{-3} mol ATP/g cdw/h.

Culture parameters like the specific CO_2 evolution rate are then determined by adding the biomass and energy mode contributions:

$$q_{CO_2}^{200} = q_{CO_2}^{x,200} + q_{CO_2}^{ATP,200} = s_1^{x,200} m_{CO_2}^{x} + s_1^{ATP,200} m_{CO_2}^{ATP} \tag{A.5}$$

Substituting the numbers from above gives the specific CO_2 evolution rate:

$$q_{CO_2}^{200} = \left(\frac{1.22 \times 10^{-7}}{\text{g cdwhr}}\right) \cdot 15078\,\text{mole}\,CO_2 + \left(\frac{1.01 \times 10^{-3}}{\text{g cdw hr}}\right) \cdot 6\,\text{mole}\,CO_2 = \frac{7.9 \times 10^{-3}\,\text{mole}\ CO_2}{\text{g cdw hr}} \tag{A.6}$$

The same procedure described here was done for all of the considered doubling times by pasting the appropriate mode

vectors into an MS Excel spreadsheet template. Adjustments in parameters like maintenance energy expenditures can be made without additional computer simulations by simply modifying the magnitude of the scaling factor. The identified pathways would not change. The biomass mode scaling factor does not change with maintenance energy requirements.

Correction: A reaction was incorrectly labeled in Appendix A.2 in Carlson and Srienc, 2004. All occurences of reaction R83 in Appendix A.2 should read reaction R82.

References

Abdel-Hamid AM, Attwood MM, Guest JR. 2001. Pyruvate oxidase contributes to the aerobic growth efficiency of *Escherichia coli*. Microbiology 147:1483–1498.

Alam KY, Clark DP. 1989. Anaerobic fermentation balance of *Escherichia coli* as observed by in vivo nuclear magnetic resonance spectroscopy. J Bacteriol 171:6213–6217.

Alexeeva S, de Kort B, Sawers G, Hellingwerf KJ, de Mattos MJ. 2000. Effects of limited aeration and of the ArcAB system on intermediary pyruvate catabolism in *Escherichia coli*. J Bacteriol 182:4934–4940.

Andersen KB, von Meyenburg K. 1980. Are growth rates of *Escherichia* in batch culture limited by respiration? J Bacteriol 144:114–123.

Bajpai R. 1987. Control of bacterial fermentations. Ann NY Acad Sci 506: 446–458.

Bauer S, Ziv E. 1976. Dense growth of aerobic bacteria in a bench-scale fermenter. Biotechnol Bioeng 28:81–94.

Bauer KA, Ben-Bassat A, Dawson M, de la Puente VT, Neway JO. 1990. Improved expression of human interleukin-2 in high-cell-density fermentor cultures of *Escherichia coli* K-12 by a phosphotransacetylase mutant. Appl Environ Microbiol 56:1296–1302.

Belaich A, Belaich JP. 1976. Microcalorimetric study of the anaerobic growth of *Escherichia coli*: growth thermograms in a synthetic medium. J Bacteriol 125:14–18.

Berrios-Rivera SJ, Yang YT, Bennett GN, San KY. 2000. Effect of glucose analog supplementation on metabolic flux distribution in anaerobic chemostat cultures of *Escherichia coli*. Metabol Eng 2:149–154.

Blackwood AC, Neish AC, Ledingham GA. 1956. Dissimilation of glucose at controlled pH values by pigmented and non-pigmented strains of *Escherichia coli*. J Bacteriol 72:497–499.

Bratbak G, Dundas I. 1984. Bacterial dry matter content and biomass estimations. Appl Environ Microbiol 48:755–757.

Burgard AP, Maranas CD. 2001. Probing the performance limits of the *Escherichia coli* metabolic network subject to gene additions or deletions. Biotechnol Bioeng 74:364–375.

Calhoun MW, Oden KL, Gennis RB, Teixeira de Mattos MJ, Neijssel OM. 1993. Energetic efficiency of *Escherichia coli*: effects of mutations in components of the aerobic respiratory chain. J Bacteriol 175: 3020–3025.

Carlson R, Srienc F. 2004. Fundamental *Escherichia coli* biochemical pathways for biomass and energy production: identification of reactions. Biotechnol Bioeng 85:1–19.

Chang DE, Jung HC, Rhee JS, Pan JG. 1999. Homofermentative production of D- or L-lactate in metabolically engineered *Escherichia coli* RR1. Appl Environ Microbiol 65:1384–1389.

Chesbro W, Evans T, Eifert R. 1979. Very slow growth of *Escherichia coli*. J Bacteriol 139:625–638.

Contiero J, Beatty C, Kumari S, DeSanti CL, Strohl WR, Wolfe A. 2000. Effects of mutations in acetate metabolism on high-cell-density growth of *Escherichia coli*. J Ind Microbiol Biotechnol 24:421–430.

de Graf MR, Alexeeva S, Snoep JL, Teixeira de Mattos MJ. 1999. The steady-state internal redox state (NADH/NAD) reflects the external redox state and is correlated with catabolic adaptation in *Escherichia coli*. J Bacteriol 181:2351–2357.

Diaz-Ricci JC, Regan L, Bailey JE. 1991. Effect of alteration of the acetic acid synthesis pathway on the fermentation pattern of *Escherichia coli*. Biotechnol Bioeng 38:1318–1324.

Edwards JS, Palsson BO. 2000. Robustness analysis of the *Escherichia coli* metabolic network. Biotechnol Prog 16:927–939.

Edwards JS, Ibarra RU, Palsson BO. 2001. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. Nat Biotechnol 19:125–130.

Edwards JS, Ramakrishna R, Palsson BO. 2002. Characterizing the metabolic phenotype: a phenotype phase plane analysis. Biotechnol Bioeng 77:27–36.

Emmerling M, Dauner M, Ponti A, Fiaux J, Hochuli M, Szyperski T, Wuthrich K, Bailey JE, Sauer U. 2002. Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. J Bacteriol 184:152–164.

Farmer IS, Jones CW. 1976. The effect of temperature on the molar growth yield and maintenance requirement of *Escherichia coli* W during aerobic growth in continuous culture. FEBS Lett 67:359–363.

Hempfling WP, Mainzer SE. 1975. Effects of varying the carbon source limiting growth on yield and maintenance characteristics of *Escherichia coli* in continuous culture. J Bacteriol 123:1076–1087.

Hong SH, Lee SY. 2002. Importance of redox balance on the production of succinic acid by metabolically engineered *Escherichia coli*. Appl Microbiol Biotechnol 58:286–290.

Kacser H, Burns JA. 1995. The control of flux. Biochem Soc Trans 23: 341–366.

Klamt S, Stelling J. 2003. Two approaches for metabolic pathway analysis? Trends Biotechnol 21:64–69.

Lee S, Phalakornkule C, Domach MM, Grossmann IE. 2000. Recursive MILP model for finding all the alternate optima in LP models for metabolic networks. Comput Chem Eng 24:711–716.

Lee SY, Papoutsakis ET. 1999. Metabolic engineering. Marcel Dekker Inc., New York.

Lendenmann U. 1994. Growth kinetics of *Escherichia coli* with mixtures of sugars. Swiss Federal Institute of Technology (ETH).

Lin HY, Mathiszik B, Xu B, Enfors SO, Neubauer P. 2001. Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations of *Escherichia coli*. Biotechnol Bioeng 73:347–357.

Luria SE. 1960. The bacterial protoplasm: composition and organization. In: Gunsalus IC, Stanier RY, editors. The bacteria. New York: Academic Press. p 1–34.

Natarajan A, Srienc F. 1999. Dynamics of glucose uptake by single *Escherichia coli* cells. Metab Eng 1:320–333.

Natarajan A, Srienc F. 2000. Glucose uptake rates of single *E. coli* cells grown in glucose-limited chemostat cultures. J Microbiol Methods 42: 87–96.

Neijssel OM, Hardy GP, Lansbergen JC, Tempest DW, O'Brien RW. 1980. Influence of growth environment on the phosphoenolpyruvate: glucose phosphotransferase activities of *Escherichia coli* and Klebsiella aerogenes: a comparative study. Arch Microbiol 125: 175–179.

Nimmo HG. 1987. The tricarboxylic acid cycle and anaplerotic reactions. In: Neidhardt FC, editor. *Escherichia coli* and *Salmonella typhimurium*. Washington, DC: American Society for Microbiology. p 156–169.

Paalme T, Elken R, Kahru A, Vanatalu K, Vilu R. 1997. The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach. Antonie Van Leeuwenhoek 71:217–230.

Papin JA, Price ND, Edwards JS, Palsson B. 2002. The genome-scale metabolic extreme pathway structure in *Haemophilus influenzae* shows significant network redundancy. J Theoret Biol 215:67–82.

Pfeiffer T, Sanchez-Valdenebro I, Nuno JC, Montero F, Schuster S. 1999. METATOOL: for studying metabolic networks. Bioinformatics (Oxford) 15:251–257.

Phalakornkule C, Lee S, Zhu T, Koepsel R, Ataai MM, Grossmann IE, Domach MM. 2001. A MILP-based flux alternative generation and NMR experimental design strategy for metabolic engineering. Metab Eng 3:124–137.

Riondet C, Cachon R, Wache Y, Alcaraz G, Divies C. 2000. Extracellular

oxidoreduction potential modifies carbon and electron flow in *Escherichia coli*. J Bacteriol 182:620–626.

Roels JA. 1983. Energetics and kinetics in biotechnology. Amsterdam: Elsevier Biomedical.

Sauer U, Lasko DR, Fiaux J, Hochuli M, Glaser R, Szyperski T, Wuthrich K, Bailey JE. 1999. Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism. J Bacteriol 181:6679–6688.

Savinell JM, Palsson BO. 1992. Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism. J Theor Biol 154:421–454.

Schilling CH, Schuster S, Palsson BO, Heinrich R. 1999. Metabolic pathway analysis: basic concepts and scientific applications in the postgenomic era. Biotechnol Progr 15:296–303.

Schultz KL, Lipe RS. 1964. Relationship between substrate concentration, growth rate, and respiration rate of *Escherichia coli* in continuous culture. Arch Mikrobiol 48:1–20.

Schuster S, Dandekar T, Fell D. 1999. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. TIBTECH. 17:53–60.

Schuster S, Hilgetag C, Fell D. 1994. Detecting elementary modes of functioning in metabolic networks. Modern Trends BioThermoKinetics 3:103–105.

Schuster S, Hilgetag C, Woods JH, Fell DA. 2002. Reaction routes in biochemical reaction systems: Algebraic properties, validated calculation procedure and example from nucleotide metabolism. J Math Biol 45:153–181.

Schuster S, Fell DA, Dandekar T. 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nat Biotechnol 18:326–332.

Schuster R, Schuster S. 1993. Refined algorithm and computer program for calculating all non-negative fluxes admissible in steady states of biochemical reaction systems with or without some flux rates fixed. Comput Appl BioSci 9:79–85.

Segre D, Vitkup D, Church GM. 2002. Analysis of optimality in natural and perturbed metabolic networks. Proc Natl Acad Sci USA 99: 15112–15117.

Senn H, Lendenmann U, Snozzi M, Hamer G, Egli T. 1994. The growth of *Escherichia coli* in glucose-limited chemostat cultures: a re-examination of the kinetics. Biochim Biophys Acta 1201:424–436.

Snoep JL, de Graef MR, Westphal AH, de Kok A, Teixeira de Mattos MJ, Neijssel OM. 1993. Differences in sensitivity to NADH of purified pyruvate dehydrogenase complexes of *Enterococcus faecalis*, *Lactococcus lactis*, *Azotobacter vinelandii* and *Escherichia coli*: implications for their activity in vivo. FEMS Microbiol Lett 114: 279–283.

Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED. 2002. Metabolic network structure determines key aspects of functionality and regulation. Nature 420:190–193.

Stephanopoulos G, Aristidou A, Nielsen J. 1998. Metabolic Engineering: Principles and Methodologies. Academic Press, San Diego.

Taylor NB. 2003. Chemical analysis of the T2 bacteriophage and its host *Escherichia coli* (strain B). J Biol Chem 271.

Tempest DW, Neijssel OM. 1987. Growth yield and energy distribution. In: Neidhardt FC, editor. *Escherichia coli* and *Salmonella typhimurium*. Washington, DC: American Society for Microbiology. p 797–806.

Varma A, Palsson BO. 1994. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. Appl Environ Microbiol 60:3724–3731.

Varma A, Boesch BW, Palsson BO. 1993. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. Appl Environ Microbiol 59:2465–2473.

Varma A, Palsson B. 1993. Metabolic capacities of *Escherichia coli* I. Synthesis of biosynthetic precursors and cofactors. J Theor Biol 165: 477–502.

Wiback SJ, Palsson BO. 2002. Extreme pathway analysis of human red blood cell metabolism. Biophys J 83:808–818.

Wiechert W. 2001. 13C metabolic flux analysis. Metabol Eng 3:195–206.