Fundamental *Escherichia coli* Biochemical Pathways for Biomass and Energy Production: Identification of Reactions

Ross Carlson, Friedrich Srienc

Department of Chemical Engineering and Materials Science, and BioTechnology Institute, 240 Gortner Laboratory, 1479 Gortner Avenue, University of Minnesota, Minneapolis/St. Paul, Minnesota, 55455/55108; telephone: 612-624-9776; fax: 612-625-1700; e-mail: fried@cbs.umn.edu

Received 12 February 2003; accepted 15 July 2003

Published online 28 October 2003 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/bit.10812

Abstract: Cells grow by oxidizing nutrients using a complex network of biochemical reactions. During this process new biological material is produced along with energy used for maintaining cellular organization. Because the metabolic network is highly branched, these tasks can be accomplished using a wide variety of unique reaction sequences. However, evolutionary pressures under carbon-limited growth conditions likely select organisms that utilize highly efficient pathways. Using elementary-mode analysis, we demonstrate that the metabolism of the bacterium *Escherichia coli* contains four unique pathways that most efficiently convert glucose and oxygen into new cells and maintenance energy under any level of oxygen limitation. Observed regulatory patterns and experimental findings suggest growing cells use these highly efficient pathways. It is predicted that five knockout mutations generate a strain that supports growth using only the most efficient reaction sequence. The analysis approach should be generally useful for predicting metabolic capabilities and efficient network designs based on only genomic information. © 2004 Wiley Periodicals, Inc.

Keywords: elementary-mode analysis; bioinformatics; cell growth; metabolism

INTRODUCTION

For more than a decade, the metabolism of *Escherichia coli* has served as a testing ground for network analysis methods. The studies of this metabolic network have been facilitated by the extensive availability of *E. coli* biochemical data. The theoretical results have provided valuable insight into the intricacies of the functioning of the metabolic network. Moreover, they have been fundamental in shaping general methodologies for studying other metabolic networks (for recent reviews, see Lee and Papoutsakis, 1999 and Schilling et al., 1999). The studies have been increasingly refined by either including more network detail or by applying more advanced and improved analysis methods. Unfortunately not all studies have been carried out on the same basis. For instance, most of the studies simplified the problem by neglecting the observation that biomass composition changes with specific growth rate even though it has been shown that this effect can be described on the basis of metabolic fluxes (Pramanik and Keasling, 1997). In addition, a simplified reaction set is often used to demonstrate the principles of an analysis method (Schilling et al., 2000a).

Essentially all studies, with rare exceptions (Liao et al., 1996; Stelling et al., 2002), have used linear programming (LP) techniques to analyze this network (Burgard and Maranas, 2001; Edwards and Palsson, 2000a, b; Varma and Palsson, 1993a, 1993b; Varma et al., 1993) or extreme pathway analysis combined with linear programming (Schilling et al., 2000a). Two exceptions used elementary mode analysis (Dandekar et al., 1999; Pfeiffer et al., 1999; Schuster and Schuster, 1993; Schuster et al., 1994, 2000, 2002) to successfully guide the metabolic engineering of aromatic pathways in *E. coli* (Liao et al., 1996) and, recently, to study transcriptional effects of growth on different substrates (Stelling et al., 2002). While this last report provided an excellent account on the analysis of a large and detailed reaction network, it did not consider two central metabolism reactions that are critical for obtaining the results described in this article. This further demonstrates that the results of any methodology are extremely sensitive to the utilized metabolic network model.

A slight variation of elementary mode analysis is extreme pathway analysis which in a previous study has been combined with linear programming (Schilling et al., 2000a). Extreme pathway analysis identifies a minimum set of generating vectors for a biochemical network's convex operating space. However, the set of extreme pathways represents a subset of elementary modes and can be difficult to interpret because the technique may not identify all functionally significant pathways (Schuster et al., 2002) (Klamt and Stelling, 2003).

Correspondence to: Friedrich Srienc

Contract grant sponsors: National Science Foundation; National Institutes of Health

Linear programming requires the formulation of an objective function for which a single network flux distribution is typically found, although a recently developed method can provide a complete set of possible fluxes for a given objective function (Lee at al., 2000; Phalakornkule et al., 2001). The obtained flux distributions are therefore very dependent on how the problem is formulated. This is in principal contrast to elementary mode analysis that identifies all pathway possibilities inherent to a metabolic network without any qualification in the problem formulation. The result of elementary mode analysis is a large set of fundamental pathways that can be sorted according to criteria similar to LP optimization functions. An attractive property of the set of elementary modes is that it describes all possible, nondivisible pathways inherent to the network. In contrast, an attractive property of LP techniques is that the single desired flux distribution can be found with much less computational effort. Of course, by repeatedly applying the LP algorithm, one can systematically map all network possibilities, an approach applied in so-called phenotypic phase plane analysis to delineate a network's metabolic capabilities (Edwards et al., 2002). This is principally different from having available a complete set of metabolism-defining building blocks. Because all pathway possibilities are explicitly defined in the elementary mode set, it contains all feasible flux distributions that can be found using LP objective functions. The task is then to sort the complete set of building blocks and find the pathways of interest.

This work presents a new and efficient method for sorting these pathways according to efficiency. This has not been done before with the *E. coli* network or, to our knowledge, with any other metabolic network subjected to elementary mode analysis (Carlson et al., 2002; van Dien and Lidstrom, 2002). A demonstration of the capabilities of the presented method has been significantly facilitated by the availability of published experimental data and published theoretical results obtained using linear programming methods, which are both utilized for comparison purposes. The analysis takes into account the variation of cell composition as a function of specific growth rate, and it considers glucose-limited growth over the entire range from completely anaerobic to completely aerobic conditions. The Appendix provides catalogued data of the optimal functioning of the metabolism over this entire range of growth conditions. It should provide, therefore, a useful resource since the most efficient metabolic flux distributions of *E. coli* can be directly read from these data for any studied condition or growth rate without further elementary mode analysis or linear programming effort. The presented analysis approach of all pathway possibilities represents a generalized technique that identifies the most efficient strategy for utilizing the available metabolic network to deal with varying degrees of culturing stress. It can be used to predict metabolic capabilities of an organism under a range of conditions and to design optimized recombinant host systems.

METHODS

Construction of Biochemical Network Model

The network model was constructed to represent *E. coli* growing on glucose minimal media (Fig. 1; for general references of an *E. coli* network, see Edwards and Palsson, 2000a and Neidhardt, 1987; 1996). 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'' (see Appendix A.1). Of the 44 reactions included in the network, 18 are considered reversible with the remaining 26 classified as irreversible. Glucose is the sole energy source while glucose and assimilated CO_2 (via PEP carboxylase) are both potential carbon sources. The model includes as external metabolites the more common by-products of *E. coli* metabolism that are often linked to oxygen-limited and anaerobic growth. Under certain conditions, these metabolites can also serve as substrates. For instance, *E. coli* can grow on acetate. However, for this study the uptake reactions associated with the metabolism of these metabolites were inactivated. It would be possible to study the role of these metabolites as substrates by changing the model's transport reactions from irreversible to reversible.

Escherichia coli possesses at least two transhydrogenase genes. The gene *pnt* encodes a membrane associated enzyme (Bragg et al., 1972) while *udhA* encodes a soluble enzyme (Boonstra et al., 1999). Transhydrogenase activity has been found in cultures grown on minimal media (Bragg et al., 1972). Since the model represents *E. coli* growth on glucose minimal media, it was assumed that a transhydrogenase activity would be present. For purposes of the model, NADH and NADPH are considered to be equivalent. Energy requirements of the membrane-associated transhydrogenase activity are assumed to be part of the maintenance energy.

The glyoxylate shunt is repressed in *E. coli* during growth on glucose (Kornberg, 1966a, 1966b; Holms and Bennett, 1971) and therefore left out of the model.

The reactions listed in the model are not necessarily the work of a single enzyme. In some cases the model was simplified by grouping nonbranching sets of reactions into a single term. For instance, R83 (see Appendix A.1) accounts for the transfer of reducing equivalents from NADH to the menaquinone pool followed by the reduction of FAD to form FADH (Lin and Kuritzkes, 1987).

The P/O numbers representing the efficiency of ATP production from NADH and FADH are assumed to be 2 and 1, respectively, based on information from Gennis and Stewart (1996) and Uden and Bongaerts (1997). The oxygen external metabolite (OXY ext, see Appendix A.1) in the model represents one oxygen atom or 0.5 mole of O_2 so the appropriate considerations have to be made when analyzing the model's output file. All presented results involving oxygen have been adjusted to represent dioxy-

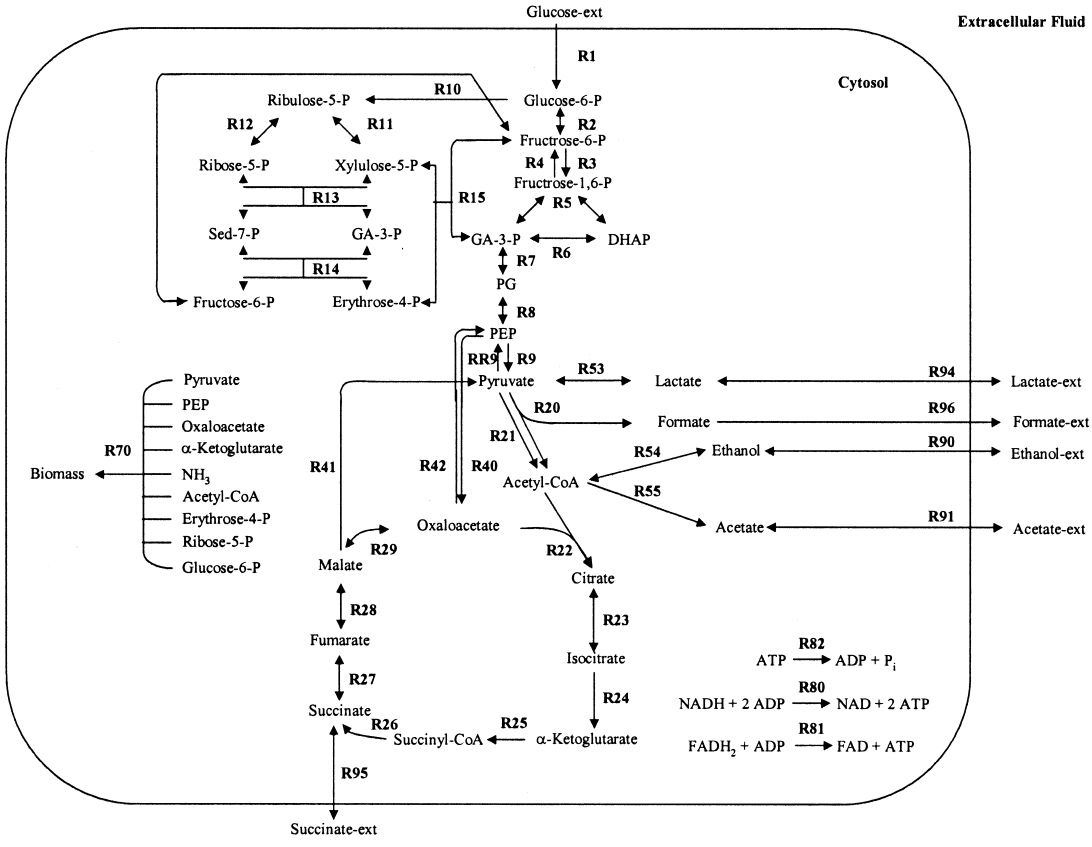

Figure 1. Map of the *E. coli* biochemical network. The reaction stoichiometry is listed separately in Appendix A. 1. The growth rate dependent coefficients for biomass production (reaction R70) are tabulated in Table II. The cofactor requirements are not shown.

gen, O_2. An external ATP metabolite is included (ATP main, see Appendix A.1) as a means to account for ATPase activity and maintenance energy requirements.

Construction of Biomass Terms

The metabolic requirements for biomass production are modelled using theory developed by Ingraham et al. (1983). This method accounts for biomass production by acknowledging the metabolite drain from the central metabolic pathways. This approach has been applied previously (Vallino and Stephanopoulos, 1990; Varma and Palsson, 1993b). Furthermore, the changing biomass composition as a function of specific growth rate has also been taken into account in the present study (Table I). Ingraham et al. pointed out that essentially all major biosynthetic components are created from 11 central metabolism precursor molecules along with ATP, NADH, NADPH, NH_3, sulphur, and single carbon units. The biomass terms consider fructose-6-P to be equivalent to gluctose-6-P due to the reversible activity of fructose-6-P isomerase. Likewise 3-phosphoglycerate is considered to be equivalent to phosphoenolpyruvate (PEP) with phosphoglycerate mutase and phosphoglycerate enolase catalyzing the reversible reactions. The metabolite contributions to the biomass terms were normalized with respect to glucose-6-P (see below). To avoid a very small triose phosphate term, the contribution of this metabolite to biomass formation was modeled as PEP using the following conversion stoichiometry: triose phosphate + NAD + ADP = PEP + ATP + NADH. These simplifications result in the biomass terms being comprised of eight central metabolic pathway intermediates (Table II). This method of accounting for the biomass production considers the drain from the central metabolism but simplifies the model by removing the need to include

Table I. Macromolecular composition of *E. coli* at different growth rates. The ''other'' category includes such components as lipopolysaccharides, peptidoglycan, polysaccharides, and lipids. Data was obtained primarily from Dennis and Bremer (1974), Bremer and Dennis (1987), Churchwald et al. (1982), and Pramanik and Keasling (1997).

Growth Rate Doubling Time (min)	Mass Fraction			
	Protein	DNA	RNA	Other
200	78	6	10	6
100	70.5	5.5	14	10
80	66.5	4.5	15	14
60	61	3	16	20
50	58	3	17	22
40	55	2	18	25
30	52	2	20	26

Table II. Stoichiometric coefficients for biomass generating reactions (reaction R70). The change in the coefficients accounts for the specific growth rate dependent biomass compositions. The specific growth rate is listed as a doubling time[a].

Doubling Time (min)	Glc-6-P	Rib-5-P	Ery-4-P	PEP	Pyr	AcCoA	α-Kg	Oxalo	CO_2	NH4	NADH	ATP	Total Carbon per Term
200	4	46	31	156	237	72	86	139	−35	731	856	2921	2652
100	4	32	17	91	129	55	47	78	−15	424	487	1674	1554
80	4	22	11	64	87	49	32	53	−10	291	347	1153	1091
60	4	16	7	45	58	44	22	36	−6	198	249	786	777
50	4	14	6	39	50	43	19	31	−4	171	218	678	673
40	4	13	5	34	41	42	16	26	−3	148	192	584	498
30	4	13	5	32	38	41	14	24	−2	139	178	547	471

[a]Glc-6-P = glucose-6-P, Rib-5-P = ribulose-5-P, Ery-4-P = erythrose-4-P, PEP = phosphoenolpyruvate, Pyr = pyruvate, AcCoA = acetyl-CoA, α-Kg = α-ketoglutarate, Oxalo = oxaloacetate. ''Total Carbon per Term'' is the total number of carbon atoms per biomass term. Each growth rate dependent term represents a different mass of cells.

separate reactions of each component of the biomass such as the production of alanine or lipopolysaccharides. The biomass terms consider the substrate, energy, and redox requirements for the synthesis of the biological precursor molecules as well as the energy requirements for polymerizing the monomers into macromolecules as discussed in Ingraham et al. (1983). The term does not include maintenance energy requirements that are treated separately. After determining the molar requirements of each component for one gram of dried biomass, the terms were normalized with respect to one fourth the mmol contribution of glucose-6-P. This value was chosen to simplify the comparison of the relative contribution of each metabolite at different growth rates, to avoid significant round-off error because the METATOOL (Version 3.5.2) elementary flux analysis program only deals with integers, and to reduce memory requirements during the computational process. The effects of this round-off error were explored and found not to have a significant effect on the results (data not shown). The biomass terms for the seven analyzed growth rates are shown in Table II. Since, the biomass terms were normalized so that the coefficient for glucose-6-P is the same, each biomass term represents a different number of C moles of biomass. This C moles per biomass term is shown in the right column of Table II.

The model does not account for sulphur used in biological processes like protein synthesis nor does it account for trace elements. In addition, the model does not contain conservation relationships for water or protons. The energy required to maintain desirable levels of these components is assumed to be included in maintenance energy expenditures.

Computer Analysis

The publicly available program METATOOL_352_double was used for all elementary mode analyses. The program reads an input file containing the network and identifies all possible, unique pathways that result in a balanced system. The output file contains the results in three formats including a matrix that contains the relative fluxes through each reaction for every mode. This program is available at http://mudshark.brookes.ac.uk/sware.html or ftp://ftp.bioinf.mdc-berlin.de/Pub/metabolic/metatool/ (Pfeiffer et al., 1999; Schuster et al., 1994).

The METATOOL output matrices were analyzed by pasting them into a 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.

RESULTS

Energy-Producing Modes

Cell growth requires more than the synthesis of the appropriate macromolecules. In addition to biosynthetic molecules, cells require a supply of maintenance energy. The significance of maintenance energy was first discussed by Pirt in 1965 (Pirt, 1965) and has since become a widely accepted concept related to cell growth (Farmer and Jones, 1976; Hempfling and Mainzer, 1975; Neijssel and Tempest, 1976; Stouthamer and Bettenhaussen, 1973). The theory implies that in addition to the fluxes associated with biomass synthesis, cells maintain metabolite fluxes to generate energy used for purposes such as maintaining ion gradients, repairing existing proteins and polynucleotides, and maintaining a biochemical infrastructure that will permit a cell to rapidly respond to sudden feast-or-famine conditions (Koch, 1971).

The energy-generating pathways were identified by sorting the elementary mode analysis output file for ATP-producing modes that did not produce biomass; 205 energy-producing pathways were identified with 32 of these pathways representing anaerobic ATP generation. The most efficient aerobic mode for ATP generation in terms of ATP yield on glucose is the classic, textbook pathway that completely oxidizes glucose into CO_2 consuming six molecules of O_2 and generating 26 ATP. This mode is

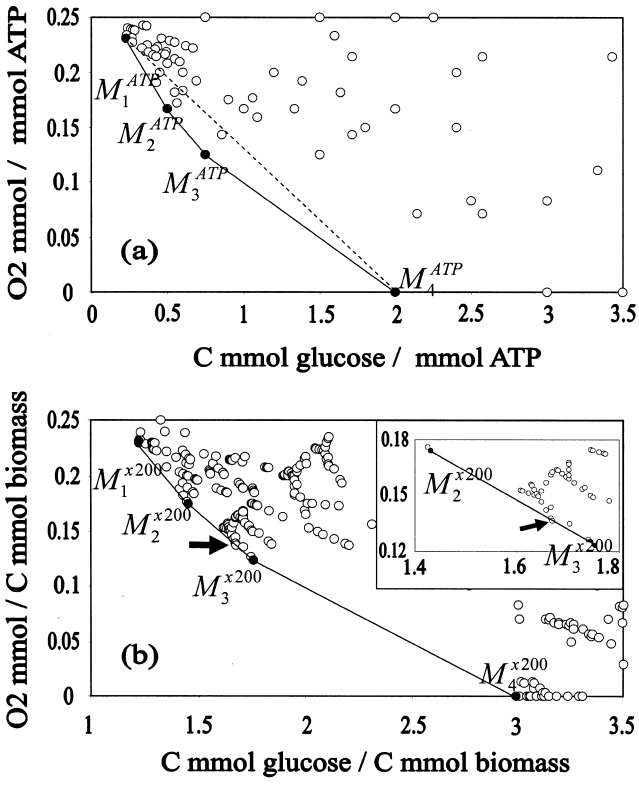

Figure 2. (**a**) Relationship between inverse ATP yield per C mmol of glucose and inverse ATP yield per mmol of O_2. Each circle represents the inverse yields of a single elementary mode. Linear combinations of the modes labeled M_1^{ATP} through M_4^{ATP} represent the most efficient means of producing ATP under conditions ranging from oxygen sufficiency to the complete lack of oxygen. More information on these modes can be found in Table III and Appendix A.2. The plot axis range was selected to highlight the most efficient data points. The dotted line demonstrates that linear combinations of M_1^{ATP} and M_4^{ATP} results in less efficient ATP production. (**b**) Plot of the inverse biomass yield per C mmol of glucose and inverse biomass yield per mmol of O_2 for a culture with a 200 minute doubling time. Each circle represents the inverse yield of a single elementary mode. The most efficient path from $M_1^{\times 200}$ to $M_4^{\times 200}$ is shown. The insert shows a close-up of the region between $M_2^{\times 200}$ and $M_3^{\times 200}$. The mode highlighted with the arrow is on the line linking modes $M_2^{\times 200}$ and $M_3^{\times 200}$. Since its behavior can be described using linear combinations of $M_2^{\times 200}$ and $M_3^{\times 200}$, it is not required for purposes of defining the system. A similar mode was found during the analysis of all the biomass models. The plot axis range was selected to highlight the most efficient data points. Points lying outside of the truncated range were not significant to the analysis.

referred to as M_1^{ATP}. The most efficient anaerobic mode produces three molecules of ATP, one molecule of acetate, one molecule of ethanol, and two molecules of formate per glucose molecule fermented. This mode is referred to as M_4^{ATP}(see Table III).

Escherichia coli cultures are often oxygen limited due to rapid growth rates and the low oxygen solubility in aqueous media. To examine ATP production under oxygen stress, the energy producing modes were first analyzed using a plot of the ATP glucose yield vs. the ATP oxygen yield (data not shown). The highest yielding modes under aerobic and anaerobic conditions are easily identified. But the plot has two drawbacks. First, while the most efficient aerobic and anaerobic modes are readily apparent it is not obvious from such a plot which modes are necessary for most efficient energy production during intermediate culturing conditions. A second drawback is that the anaerobic yield becomes infinity—which is inconvenient to represent. The two problems are easily addressed using a plot of the inverse ATP yields based on oxygen and on glucose (Fig. 2a). The yields were obtained from the stoichiometric coefficients of each energy-producing mode. Each data point in Figure 2a represents the oxygen and glucose requirements to produce one mmol of ATP for each of the 205 identified energy modes. Since each data point in the figure represents the nutrient consumption for the same amount of ATP, the modes with the lowest combined requirement for both oxygen and glucose are the most efficient. These points are located closest to the origin, along the lower left edge of the plot. Intermediate metabolic states not represented by the discrete data points can be created by linear combinations of unique modes.

The network's most efficient strategy for spanning all conditions between oxygen sufficiency (M_1^{ATP}) and the complete lack of oxygen (M_4^{ATP}) involves an envelope of line segments that links four energy modes. The envelope connects M_1^{ATP} with M_4^{ATP} by passing through two additional modes M_2^{ATP} and M_3^{ATP}. As can be seen in Figure 2a, the path defined by the line segments represents the most efficient use of the available modes. For instance, a line drawn directly between M_1^{ATP} and M_4^{ATP} would pass above metabolic states 2 and 3 (M_2^{ATP} and M_3^{ATP}) and would therefore represent a less efficient strategy since it would require more oxygen and glucose to produce the same amount of ATP. The technique is a graphical, two parameter optimization procedure which minimizes the glucose and oxygen requirements for ATP production on the basis of available elementary modes. The metabolic states found along the line segments are defined by linear combinations

Table III. Aerobic and anaerobic ATP producing modes with associated ATP carbon yield, ATP oxygen yield, and overall stoichiometry[a].

Mode:	$Y_{ATP/C}$	$Y_{ATP/O2}$	Mode Stoichiometry
$M_{ATP.1}$	4.33	4.33	glucose + 6 O_2 = 26 ATP + 6 CO_2
$M_{ATP.2}$	2	6	glucose + 2 O_2 = 12 ATP + 2 acetate + 2 CO_2
$M_{ATP.3}$	1.33	8	glucose + O_2 = 8 ATP + 2 acetate + 2 formate
$M_{ATP.4}$	0.5	anaerobic	glucose = 3 ATP + ethanol + acetate + 2 formate

[a]See Figure 4 and Appendix A.2 for the flux pattern for each mode. $Y_{ATP/C}$ = moles ATP/mole glucose carbon consumed, $Y_{ATP/O2}$ = moles ATP/mole of dioxygen consumed.

of the two elementary modes that serve as the endpoints. Biologically significant intermediate metabolic states are found only between identified modes that can contribute at a certain fraction. By definition, a meaningful fraction can only assume a value between 0 and 1. Therefore, intermediate states that are based on two contributing elementary modes must lie on a line connecting the two modes. The thick lines in Figure 4 show the reactions utilized by the four modes.

The stoichiometric coefficients of the four pathways define the network's optimal relationship between oxygen, glucose, and ATP for different culturing conditions. Figure 3a maps the ATP production rate to the oxygen and glucose fluxes within the operating space as defined by the mode stoichiometry. The slope of the ATP production isoline changes depending on the availability of oxygen or glucose. Following an ATP production isoline from left to right corresponds to a culture experiencing progressively more severe oxygen limitation with anaerobic conditions occurring when the line reaches the abscissa. It can be seen that the specific glucose uptake rate must increase if the same amount of ATP is to be produced with a lowered specific oxygen flux. The energy generating pathways do not involve synthesis of biomass. This indicates that energy can be produced independently from the biomass pathway reactions.

Biomass-Producing Modes

Because biomass composition changes with growth rate (see Table I), each growth rate case was analyzed separately taking into account the changed reaction stoichiometry of biomass synthesis. Similar to energy production, the analysis revealed a large number of possible strategies for producing biomass. For instance, the model representing a 200-min doubling time had 862 different biomass producing modes with 184 of these modes being anaerobic. As with energy production, some of these modes represent more efficient pathways than others.

As with the analysis of the energy-producing modes, a plot of the inverse biomass yields based on oxygen and on glucose (Fig. 2b) reveals the most efficient pathways. The most efficient path between oxygen sufficiency and anaerobic biomass production passes through four modes. Connecting these modes, $M_1^{\times 200}$ through $M_4^{\times 200}$, with straight line segments creates an envelope that represents the most efficient use of the available biochemical reactions to produce biomass under any level of oxygenation. Any intermediate metabolic state found along the line segments is defined by linear combinations of the two end point modes. The contribution from each mode is restricted by the same considerations as discussed with the energy modes. The reactions utilized by the four modes are shown in Figure 4.

The data given in Figure 2b is for a culture with a 200-min doubling time. The same procedure was applied for all of the tested growth rates. The same trends were seen in each case (see Table IV and Appendix A.3). Four modes were

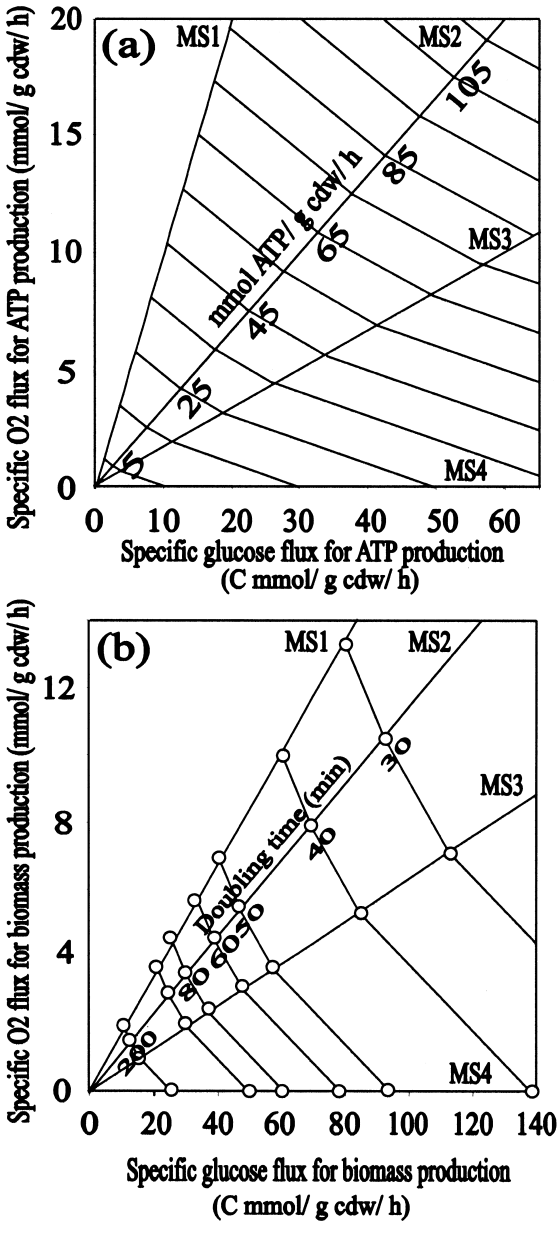

Figure 3. **(a)** Relationship between specific oxygen uptake rate, specific glucose uptake rate, and specific ATP production rate for the *E. coli* network. The specific ATP production rate axis is labeled on top of the line for the energy producing mode M_2^{ATP}. The energy producing modes M_1^{ATP} through M_4^{ATP} define the optimal relationship between carbon and oxygen flux for energy production. **(b)** Relationship between specific oxygen uptake rate, specific glucose-uptake rate and specific growth rate for *E. coli* biomass production. The specific growth rate axis is labeled on top of the line for metabolic state 2 (MS2). The metabolic state lines are linear regression fits of the data points shown in the plot as circles. The line fit r^2 values were better than or equal to 0.995 for all three lines. This chart is solely for the production of biomass. It does not account for the maintenance energy requirements. The steady-state specific glucose uptake requirements for biomass production were determined using the following conversion: $q_{glc} = \mu(0.48 \text{ g carbon/g cdw})(1 \text{ C mol}/12 \text{ g carbon})(1/Y_{x/glc,i})$. The biomass carbon yield ($Y_{x/glc,i}$) for each metabolic state (i) and for each growth rate can be found in Appendix A.3. The cells are assumed to be 48% carbon by mass. This value is not believed to be a function of growth rate (Hempfling and Mainzer, 1975). cdw = cell dry weight.

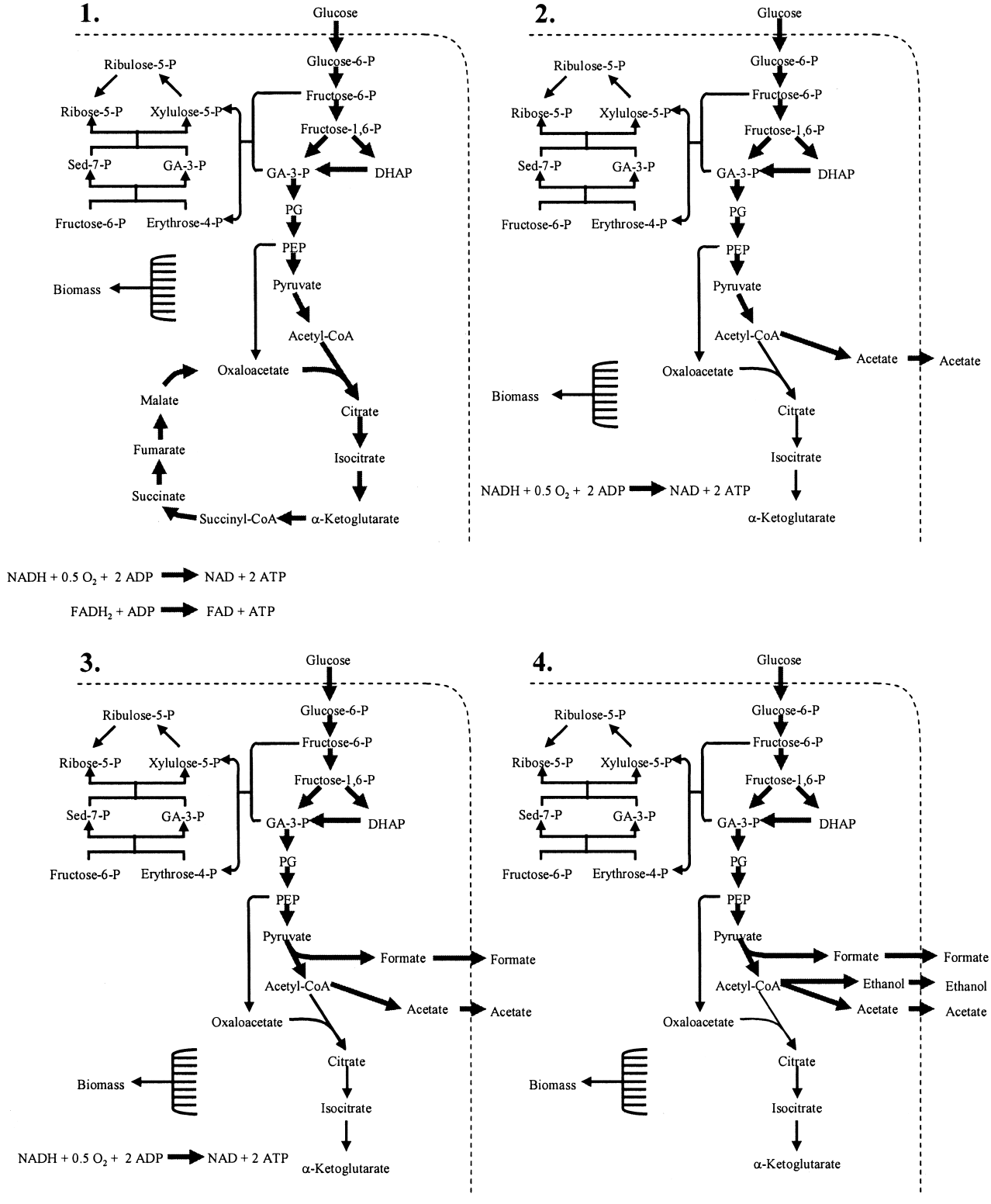

Figure 4. Map of the energy (thick lines) and biomass (thick and thin lines) producing modes for metabolic states 1 through 4. The modes for energy and biomass production can be superimposed on top of each other. The four metabolic states represent the most efficient strategies for any level of oxygen limitation. The overall stoichiometry of the modes is listed in Table III and Appendix A.3. The relative rates through each reaction can be found in Appendix A.2 and A.4.

Table IV. Biomass-producing modes for different specific growth rates.

Doubling Time[a] (min)	Total Number of Biomass-Producing Modes[b]	Most Efficient Aerobic Biomass Mode Yield Based on Carbon	Most Efficient Aerobic Biomass Mode Yield Based on Oxygen	Number of Anaerobic Biomass-Producing Modes	Most Efficient Anaerobic Biomass Mode Yield Based On carbon
200	862	0.821	4.36	184	0.335
100	860	0.823	4.45	184	0.341
80	860	0.822	4.56	184	0.344
60	850	0.825	4.81	182	0.355
50	856	0.813	4.86	184	0.354
40	859	0.687	4.18	188	0.302
30	859	0.690	4.21	188	0.304

[a]The specific growth rate is given as a doubling time in minutes.
[b]In addition to the biomass-producing modes listed, 829 modes are found that do not produce biomass.

required to create the most efficient envelope covering culturing conditions between oxygen sufficiency and the complete lack of oxygen. It should be noted that in each case the corresponding four metabolic states utilized the exact same combination of reactions. However, the ratio of fluxes through these reaction steps varied in accordance with the different cell compositions found at different growth rates. The most efficient relationship between specific oxygen uptake rates, specific glucose uptake rates, and the specific growth rate is shown in Figure 3b. The values in this plot are for biomass production only and do not include maintenance energy requirements. The predicted nutrient requirements for each metabolic state, regardless of growth rate and the associated differences in biomass composition, are closely fit by a straight line.

Identification of the four modes permits a detailed examination of the network's performance. For instance, analysis of modes 3 and 4 predicts the necessity of the formate hydrogen-lyase enzyme complex for cultures to grow most efficiently under microaerobic or anaerobic conditions. Mode 4 requires 104 moles of CO_2 as a substrate per 200-min doubling time biomass term which under the considered steady-state conditions must come from the cleavage of formate into CO_2 and H_2 via the formate hydrogen-lyase enzyme complex (see stoichiometry in Appendix A.3). Under conditions described by modes 3 and 4, the TCA cycle is branched into an oxidative and a reductive arm (Nimmo, 1987). The only means of producing the oxaloacetate required for biomass production is through the PEP carboxylase (reaction R40; see Fig. 1 or Fig. 4) mediated assimilation of CO_2 and PEP. However, under microaerobic and anaerobic conditions (modes 3 and 4), the activity of PHDc is reduced or inhibited so isocitrate dehydrogenase (R24) is the only major CO_2 producing reaction active in the central metabolism. The oxidative branch of the pentose phosphate pathway (R10) is not required for growth on minimal media because of the active transhydrogenase system. The relative contributions of α-ketoglutarate and oxaloacetate for biomass production show that flux through isocitrate dehydrogenase will not supply enough CO_2 for all of the required oxaloacetate (see Table II). For a 200-min doubling time, there are 86 molecules of CO_2 formed from the synthesis of α-ketoglutarate while there is a requirement of 139 molecules of CO_2 for the production of oxaloacetate. Therefore, the formate hydrogen-lyase enzyme complex is likely necessary to supply the CO_2 needed for PEP carboxylase.

Growth requires the combined production of biomass and maintenance energy. The possibility of biomass and energy production being intrinsically linked into a single mode was examined. Sixteen modes produced biomass along with an excess of ATP. The most efficient of these modes in terms of combined biomass and energy production were 1–3% less efficient than a combination of a pure biomass mode (M_1^x) and a pure energy mode (M_1^{ATP}) (data not shown).

Comparison of Energy- and Biomass-Producing Pathways

The energy- and biomass-producing modes were independently found to use a combination of four modes to span all culturing conditions between oxygen sufficiency and the complete absence of oxygen. It is remarkable that each identified energy-producing mode has a corresponding counterpart for biomass production. In fact, it is clear from Figure 4 that the four energy-producing modes can operate on the basis of the corresponding biomass mode. This follows from the requirement of ATP for biomass production (see Table II). Strategies that are efficient for ATP production will also be efficient for biomass production. The reaction network is constructed in a manner that permits the superimposing of a mode that generates only ATP on top of a second elementary mode that produces biomass (see Fig. 4). During growth, the overall flux state of a cell is expected to consist of fluxes associated with both energy and biomass production. Since these fluxes coexist using the same cellular machinery, their regulation is expected to be based on the same mechanisms. For every growth rate, metabolic state 1 (MS1) has the highest biomass yield on glucose and the lowest biomass yield on oxygen of the four metabolic states

found on the envelope. Because the energy-generating mode can be overlaid onto the biomass modes, the same yield trends also apply for the corresponding energy pathway. Metabolic state 1 utilizes the TCA cycle and PEP carboxylase (R40) to produce oxaloacetate. See Table V for a summary of key enzymatic regulation. Metabolic state 2 has an intermediate biomass glucose and oxygen yield. Metabolic state 2 does not use the TCA cycle and all of the oxaloacetate required for biomass production comes from PEP carboxylase. The reduced glucose yield is due, in part, to the secretion of the partially oxidized metabolite acetate. Metabolic state 3 is similar to MS2 except pyruvate formate-lyase is utilized instead of the pyruvate dehydrogenase complex (PDHc). This mode has the lowest glucose yield of the aerobic biomass generating modes but the highest oxygen yield. The mode secretes two partially oxidized metabolites acetate and formate. Metabolic state 4 produces biomass anaerobically with substrate level phosphorylation used for all ATP requirements. Three partially oxidized metabolites—acetate, ethanol and formate—are produced.

Since the biomass and energy modes can be superimposed upon each other, a description of the cell's overall flux state can be made using combinations of the energy and biomass modes. Culturing conditions in between the four defined metabolic states are described as a linear average of the two end-point metabolic states. Since the energy and biomass modes utilize the same enzymes subject to the same regulatory signals, an intermediate metabolic state based on a linear combination of M_1^{ATP} and M_2^{ATP} is also located between $M_1^{\times 200}$ and $M_2^{\times 200}$ defined by the corresponding weighting factors for the two modes.

Observed Regulation Patterns Are Consistent With Identified Metabolic States

If a cell is to operate at a flux state described by the identified pathways, the network's regulatory mechanisms would need to be able to adapt to available nutrient conditions by selecting the appropriate mode. For instance, to shift from an oxygen-sufficient metabolic state described by M_1^{ATP} and $M_1^{\times 200}$ to an oxygen-limited state described by M_2^{ATP} and $M_2^{\times 200}$, the cyclic flux through the TCA cycle would need to be inhibited since only the oxygen sufficient mode utilizes a completely operational TCA cycle (see Fig. 4). α-Ketoglutarate is a required metabolite for biomass production (see Table II). While growing on glucose minimal media, the oxidative branch of the TCA cycle is the only route available to produce this metabolite. The most logical enzyme target to mediate the transition from MS1 to MS2, while maintaining the production of α-ketoglutarate, is α-ketoglutarate dehydrogenase (R25, see Fig. 1 and Fig. 4). Inhibition of this activity would permit the necessary production of α-ketoglutarate for biomass synthesis while at the same interrupting cyclic flux through the TCA cycle. This would cause a metabolic shift toward a flux state described by MS2. These predictions appear to be consistent with reported experimental data since α-ketoglutarate dehydrogenase is down-regulated at both a genetic and enzymatic level during oxygen limitation (Amarasingham and Davis, 1965; Thomas et al., 1972). In addition, the modes predict that a shift from MS1 to MS2 would result in acetate production. This behavior is also seen experimentally. Acetate is commonly found in rapidly growing and/or oxygen limited *E. coli* cultures (Tempest and Neijssel, 1987). Additional experimental and theoretical data related to acetate production have been previously described and mechanisms of acetate overflow in *E. coli* have been proposed (Majewski and Domach, 1990; van de Walle and Shiloach, 1998). The proposed mechanisms are consistent with the implications of a cell operating at a metabolic state described by MS1 and MS2.

To switch from a metabolic state described by M_2^{ATP} and $M_2^{\times 200}$ to the state described by M_3^{ATP} and $M_3^{\times 200}$, the pyruvate-formate lyase (PFL) enzyme (see Fig. 4) needs to be activated. Because the radical containing active site of PFL is irreversibly damaged by oxygen, it was once believed that this enzyme is active only under completely anaerobic conditions (Conradt et al., 1983; Knappe, 1987). However, this hypothesis has been challenged recently by work that reveals that the enzyme is also active under microaerobic conditions (Alexeeva et al., 2000; Fiaux et al., 1999). In addition to experimental data supporting the activity of PFL under microaerobic conditions, the expression profiles of different cytochromes during oxygen limitation have been examined (Alexeeva et al., 2000). Under microaerobic conditions, a cytochrome with a high affinity for oxygen is expressed which serves as a proposed respiratory protective mechanism. The high affinity cytochrome is thought to protect the PFL enzyme by reducing soluble cytosolic oxygen concentrations. The existence of a regulatory mechanism that permits the activity of PFL under microaerobic conditions is consistent with regulation required by MS3.

The transition between MS3 and MS4 predicted by the mode analysis involves the cessation of flux through oxidative phosphorylation (R80) and the production of ethanol. Metabolic state 4 is a completely anaerobic metabolic state therefore oxidative phosphorylation would not occur. Second, the near equimolar production of both acetate and ethanol predicted by MS4 is supported by experimental data. A sampling of literature shows that acetate and ethanol production levels during anaerobic growth are usually quite similar (Alam and Clark, 1989; Belaich and Belaich, 1976; Blackwood et al., 1956; Chesbro et al., 1979). Like MS3, MS4 requires the activity of PFL. The activity of this enzyme under anaerobic conditions is well established (Knappe, 1987). As with the other three metabolic states, the predictions from these modes are consistent with the observed behavior of *E. coli* suggesting that evolutionary pressures created a regulatory system that permits the most efficient use of the available biochemical reactions. A summary of key enzyme regulation is shown in Table V.

Table V. Summary of key enzymatic features associated with each of the four metabolic states (MS).

Metabolic State	TCA Flux[a]	PDHc/PFL[b]	Oxaloacetate[c]	By-products[d]	FHL[e]
MS1	Cyclic	PDHc	ppc, mdh	None	No
MS2	Branched	PDHc	ppc	Acetate	No
MS3	Branched	PFL	ppc	Acetate, formate	Yes
MS4	Branched	PFL	ppc	Acetate, formate, ethanol	Yes

[a]Details flux patterns through the tricarboxylic acid cycle.
[b]Details the use of either pyruvate dehydrogenase complex (PDHc) or pyruvate formate-lyase (PFL) to produce acetyl-CoA from pyruvate.
[c]Details the pathways used to produced oxaloacetate with ppc = PEP carboxylase and mdh = malate dehydrogenase.
[d]Details partially oxidized by-product secretion.
[e]Details the requirement of the formate hydrogen-lyase enzyme for optimal biomass production.
See text for more details.

Another prediction that can be experimentally verified deals with the activity of the PEP carboxylase enzyme. Every identified biomass-producing mode, regardless of growth rate, utilizes the PEP carboxylase (R40) to produce some, if not all, of the oxaloacetate required for biomass production. For instance, all 862 biomass-producing modes for the model representing the culture with the 200-min doubling time used the PEP carboxylase reaction. There does not appear to be any steady-state means of producing biomass from glucose minimal media without the use of this enzyme. This is remarkable because there are two pathways leading to the production of oxaloacetate. In addition to the PEP carboxylase route, oxaloacetate can be produced via the TCA cycle. However, the TCA cycle cannot, under steady-state conditions, supply all of the metabolite required for biomass production. This prediction is consistent with experimental data that has shown *E. coli* is unable to grow on glucose minimal media without the PEP carboxylase enzyme (Coomes et al., 1985; Kornberg, 1966a, 1966b).

Robustness of Efficient Pathways

A number of studies have addressed the property of biochemical network robustness (Edwards and Palsson, 2000b; Papin et al., 2002; Price et al., 2002; Schuster et al. 2000; Stelling et al., 2002). For instance, in Edwards and Palsson (2000b) robustness was studied using linear programming techniques by systematically reducing permitted flux through select reactions. Since elementary mode analysis identifies all possible pathways, it is ideally suited to comprehensively analyze pathway robustness. For instance, Stelling et al. (2002) reported for their model that there were more than 43,000 possible modes available to *E. coli* metabolizing the substrates glucose, acetate, glycerol, or succinate. We further develop this analysis by focusing on the robustness of highly efficient pathways for biomass and energy production under varying levels of oxygen stress. The presented technique for identifying the most efficient modes also permits the quick and efficient identification of other highly efficient modes and therefore, gauges an additional aspect of network robustness. For instance, with the 200-min doubling time biomass model, there are 24 different completely aerobic modes that have biomass carbon yields between 0.80 and 0.82. Using some published experimental observations, it is possible to remove the biologically unlikely modes from this list. Under oxygen-sufficient and glucose-limited conditions, *E. coli* produces only CO_2 as a by-product (Tempest and Neijssel, 1987). Five of the 24 modes produce by-products other than CO_2 so they were excluded from this analysis. A comparison of the relative flux patterns associated with each of the remaining 19 modes is given in Table VI. The columns associated with the 16 reactions that appear in every mode are shaded gray. These reactions are essential for the most efficient aerobic production of biomass. These reactions are likely the most significant indicators of the flux state since they must be active for efficient growth on glucose. On an individual basis, the other reactions are not required since alternative pathways can replace their activity with a similar level of efficiency. This suggests a cell has little selective pressure to tightly control the partitioning of flux around these reactions since a large number of possible variations result in an essentially equally efficient biomass production.

As in the oxygen-sufficient case, the anaerobic results exhibit network robustness. There are nine unique modes that produce biomass anaerobically with carbon yields ranging from 0.32 to 0.34 (see Table VII). Anaerobic conditions restrict the system more than aerobic conditions as can be seen by the larger fraction of the total reactions required for efficient anaerobic growth.

Both the oxygen sufficient and completely anaerobic modes show redundancy and robustness. The presented methodology permits identification of not only the most efficient pathways under oxygen sufficient and anaerobic conditions but it also identifies pathway robustness at culturing states in between these two extremes. The envelope from $M_1^{\times 200}$ to $M_4^{\times 200}$ represents the most efficient path between oxygen sufficiency and completely anaerobic states but as can be seen in Figure 2b there are a number of other modes with only slightly less efficient characteristics. These modes represent additional network robustness.

The presented case examines oxygen stress. The less efficient pathways likely exist because some of the reactions are utilized and provide an advantage for other types of culturing stress such as heat shock, osmotic stress, and ni-

Table VI. Comparison of the 19 most efficient aerobic modes for biomass production for a steady-state *E. coli* culture growing with a 200-minute doubling time. The relative fluxes were all normalized with respect to the production of 1 biomass unit (see Table II). The most efficient mode ($Y_{x/c} = 0.82$) is listed first while the least efficient of the considered modes is listed last ($Y_{x/c} = 0.80$). The columns associated with reactions used by all 19 modes are shaded gray. See text for more details.

Mode:	Glycosis								PPP[a]					PDHc	TCA Cycle			Anaplerotic Reactions			Oxidative Phosphorylation			
	R1	R2	R3	R4	R5/6	R7/8	R9	RR9	R10	R11	R12	R13/14*	R15*	R21	R22/24	R25/28	R29	R40	R41	R42	R80	R81	R82	R83
1.	539	535	494	0	494	951	31.7	0	0	−41	41	−5	−36	333	261	175	175	225	0	0	1041	175	0	0
2.	539	520	489	0	489	947	26.8	0	15	−31	46	0	−31	329	257	171	171	225	0	0	1047	171	0	0
3.	539	474	474	0	474	932	11.9	0	61.5	0	61.5	15.5	−16	314	242	156	156	225	0	0	1069	156	0	0
4.	540	437	462	0	462	921	0	0	98.8	24.8	73.9	27.9	−3.1	303	231	145	145	225	0	0	1086	145	0	0
5.	540	474	474	0	474	933	0	0	61.5	0	61.5	15.5	−16	315	243	157	145	237	12.4	0	1074	157	0	0
6.	540	536	495	0	495	954	0	0	0	−41	41	−5	−36	336	264	178	145	258	32.9	0	1053	178	0	0
7.	540	521	490	0	490	949	0	0	15	−31	46	0	−31	331	259	173	145	253	279	0	1058	173	0	0
8.	540	428	459	0	459	918	0	2.85	108	31	77	31	0	300	228	142	142	225	0	0	1091	142	0	0
9.	543	431	462	0	462	924	0	0	108	31	77	31	0	306	234	148	148	302	0	77	1120	148	0	0
10.	543	431	462	0	462	924	0	0	108	31	77	31	0	306	234	148	148	225	0	0	1120	148	77	0
11.	543	431	539	77	462	924	0	0	108	31	77	31	0	306	234	148	148	225	0	0	1120	148	0	0
12.	543	431	462	0	462	924	0	0	108	31	77	31	0	306	234	148	148	225	0	0	1043	225	0	77
13.	551	0	324	0	324	794	0	138	547	324	223	177	146	176	104	17.6	17.6	225	0	0	1346	17.6	0	0
14.	553	−61	304	0	304	776	0	158	609	365	244	198	167	158	86	0	0	225	0	0	1382	0	0	0
15.	553	487	487	0	487	959	0	158	61.5	0	61.5	15.5	−16	341	269	183	0	408	183	0	1200	183	0	0
16.	553	0	325	0	325	796	0	158	549	325	224	178	147	178	106	20.2	0	245	20.2	0	1362	20.2	0	0
17.	553	441	472	0	472	943	0	158	108	31	77	31	0	325	253	167	0	392	167	0	1215	167	0	0
18.	553	549	508	0	508	979	0	158	0	−41	41	−5	−36	361	289	203	0	428	203	0	1179	203	0	0
19.	553	534	503	0	503	974	0	158	15	−31	46	0	−31	356	264	198	0	423	198	0	1184	198	0	0

Note: The reactions for CO_2 production and NH_3 uptake are not included. See Appendix A.1 for the designation of each reaction number.

*The transketolase enzyme catalyzes both R13 and R15 which is why both columns are highlighted gray.

[a]PPP = pentose phosphate pathway.

[b]PDHc = pyruvate dehydrogenase complex.

[c]TCA = tricarboxylic acid.

Table VII. Comparison of the five most efficient anaerobic modes for biomass production for a steady-state *E. coli* culture growing with a 200-minute doubling time. The relative fluxes given below were all normalized with respect to the production of 1 biomass unit (see Table II). The most efficient mode ($Y_{x/c} = 0.34$) is listed first while the least efficient of the considered modes is listed last ($Y_{x/c} = 0.32$). The columns associated with reactions used by all five modes are shaded gray. See text for more details.

Mode:	Glycosis						PPP[a]					PFL[b]	PDHc	TCA Cycle[c]		PEP carb[d]	EtOH	Acetate	Formate	CO_2
	R1	R2	R3	R5	R6	R7/8	R9	R10	R11	R12	R13*	R14	R15*	R20	R21	R22/24	R25/28	R40	R54/90	R55/91	R96	R97
1.	1321	1317	1276	1276	−1276	2516	814	0	−41	41	−5	−5	−36	1898	0	86	0	225	873	867	1898	−139
2.	1329	1310	1279	1279	−1279	2526	817	15	−31	46	0	0	−31	1908	0	86	0	225	893	857	1908	−124
3.	1352	1286	1286	1286	−1286	2557	824	62	0	62	16	16	−16	1939	0	86	0	225	955	826	1939	−77.5
4.	1375	1263	1294	1294	−1294	2588	832	108	31	77	31	31	0	1970	0	86	0	225	1017	795	1970	−31
5.	1391	1248	1299	1299	−1299	2609	837	139	52	87	41	41	10	1991	0	86	0	225	1058	774	1991	0

Note: The reaction for NH_3 uptake is not included. See Appendix A.1 for the designation of each reaction number.

*The transketolase enzyme catalyzes both R13 and R15. Therefore, both columns are highlighted gray.

[a]PPP = pentose phosphate pathway.

[b]PFL = pyruvate formate-lyase.

[c]TCA = tricarboxylic acid.

[d]PEP carb. = PEP carboxylase.

trogen limitation, or for the metabolism of other substrates like pentose sugars.

Gene Knockout Targets for Efficient Growth on Glucose

The information in Tables VI and VII can be used to identify enzymatic knockout targets that could theoretically create a network capable of only the most efficient conversion of glucose into biomass. To create a network capable of only the highest yielding mode shown in Table VI, only three enzymatic activities need to be removed from the model. If reaction R10 which represents the oxidative branch of the pentose phosphate pathway is removed, the total number of potential biomass producing modes drops from 862 to 169. Next, if R41 which represents the enzymes of the malic enzyme(s) is removed, the number of potential biomass producing modes drops to 102. Of the remaining modes, 100 produce by-products like formate, ethanol, acetate, or lactate. Under typical oxygen sufficient conditions, the enzymatic activities associated with the production of these by-products are not active due to the native *E. coli* regulatory network (Tempest and Neijssel, 1987) so these modes can be excluded when considering only oxygen sufficient growth. This leaves two modes for biomass production. By removing the activity associated with the action of NADH dehydrogenase II (R83), the number of potential modes drops to one. With the removal of three enzymatic activities, it is theoretically possible to create a network, which under oxygen sufficient conditions can only produce biomass using the most efficient pathway.

If this analysis is extended to consider all levels of oxygen limitation, only two additional activities have to be removed to create a network capable of only the most efficient conversion of glucose into biomass. The two activities are associated with the production of lactic acid and succinate as metabolic by-products. Removal of these two activities from the network along with the three previously mentioned activities would limit the network to 12 possible biomass producing modes. Of the remaining 12 modes, seven modes utilize reactions that would normally be repressed by the native *E. coli* regulatory system and therefore represent biologically unlikely modes. For example, several of the modes use the PDHc for the anaerobic production of acetyl-CoA. Under typical anaerobic conditions, PDHc activity is repressed (Hansen and Henning, 1966) so these modes can be excluded. The five remaining modes are the identified modes for metabolic states 1–4 and the mode shown in Figure 2b to lie exactly on the line between $M_2^{\times 200}$ and $M_3^{\times 200}$. By removing five enzymatic activities, as suggested by the model, the system could be fine-tuned for efficient biomass production on glucose minimal medium under any level of oxygen limitation.

The experimental techniques associated with gene knockouts are well developed and information on strains with some of these activities removed individually has been already described (Calhoun et al., 1993; Fraenkel, 1968;

Hirsch et al., 1963; Yang et al., 1999). Complications associated with maintenance energy features like futile cycles would have to be addressed separately. The presented network model contains three futile cycles. These futile cycles were identified by analyzing the network with the restriction that ATP is the only permitted consumable external metabolite. The cycles exist between phosphofructokinase and fructose 2,6-bisphosphatase (R3 and R4), PEP carboxylase and PEP carboxykinase (R40 and R42), as well as between PEP synthase, malate dehydrogenase, PEP carboxylase, and malic enzyme (RR9, R29, R40, and R41).

DISCUSSION

Elementary mode analysis identifies all possible metabolite fluxes on the basis of nondivisible pathways for a given network. An efficient methodology has been developed which is capable of mining a network's complete set of possible metabolic fluxes for pathways of interest. In this study it has been shown that the most efficient metabolite fluxes for energy or biomass production anywhere in the defined operating space can be described by a linear combination of no more than two of the identified basic metabolic pathways. The analysis reveals that the network is capable of coping with culturing stresses like oxygen limitation by changing only a limited number of fluxes. The pathways optimize energy or biomass production by lowering one yield like the carbon yield to improve another associated yield like the oxygen yield to most efficiently utilize the available resources.

The four identified metabolic states are similar to results previously described using linear programming (Varma et al., 1993). In Varma et al., the equivalent of five ''metabolic states'' were identified. The difference in the number of identified states is due to the use of different P/O numbers. Varma et al. used a P/O number of 1.33 for NADH which reflects the stoichiometry of NADH dehydrogenase II while in the current study, we use a P/O number of 2 based on the stoichiometry of NADH dehydrogenase I (Gennis and Stewart, 1996; Unden and Bongaerts, 1997). Acetate secretion data is presented in a parametric sensitivity analysis as support for a P/O number of 1.33 (Varma and Palsson, 1995). However, we are able to predict acetate secretion profiles remarkably similar to Varma and Palsson's experimental data using a P/O number of 2 and a different method for determining maintenance energy requirements (Carlson and Srienc, submitted). One goal of the presented study is to identify the most efficient pathways available to an *E. coli* network. NADH dehydrogenase I represents, at least theoretically, a more efficient means of generating a proton gradient than NADH dehydrogenase II. We consider the action of the soluble transhydrogenase (*udhA;* Boonstra et al., 1999) because, at least theoretically, it represents a more efficient reaction than the membrane-bound enzyme since it does not require energy from the proton gradient.

The presented elementary mode analysis study provides an alternative method of generating phenotypic phase planes. Previously described methods (Edwards et al., 2002) explored regions of the phase plane by determining shadow price values as a function of glucose and oxygen fluxes. Our presented methodology constructs the phase plane by identifying the most efficient, unique, nondivisible metabolic pathways available to span a range of culturing conditions. This can be done with a single computer simulation and without subjective optimization criteria. Conditions between these distinct metabolic states are then most efficiently spanned by straight line segments that reflect intermediate metabolic states constructed from non-negative linear combinations of the two endpoints. The constructed phase planes can be interpreted on the basis of the stoichiometry of distinct metabolic states as opposed to the mathematical concept of a shadow price. The various phases of the plot are the result of the metabolic network maximizing efficiency for the available substrate fluxes by exchanging the yields of one metabolic state for another. For instance, a metabolic state with a high ATP per glucose yield and a low ATP per oxygen yield can be exchanged through enzyme regulation for a metabolic state which has an improved ATP per oxygen yield but a lowered ATP per glucose yield. The trade-off in yields results in the most efficient consumption of the available oxygen and glucose fluxes. The current study also takes into account growth rate dependent changes in biomass composition; this permits a more detailed description of intracellular fluxes. There are very different relative intracellular fluxes depending on the culture's growth rate. For instance, with a 200-min growth rate, biomass production requires 237 molecules of pyruvate for every 4 molecules of glucose-6-P while with a 30-min doubling time, biomass production requires only 39 molecules of pyruvate for each 4 molecules of glucose-6-P (see Table II).

Escherichia coli networks have been studied using a combination of linear optimization, phenotypic phase plane, and extreme pathway analysis (Schilling et al., 2000a). In this study they presented an ''example'' metabolic network that did not consider some critical reactions found in the central metabolism. For instance, the pyruvate formate-lyase enzyme was not considered, that limited the scope of their study to conditions that in this study would correspond to MS1 and MS2. MS3 and MS4 require the activity of the pyruvate formate lyase enzyme complex. In our presented study we used elementary mode analysis to first identify all unique, energy- and biomass-producing pathways including those pathways that utilize the pyruvate formate-lyase enzyme. Then an efficient, relatively simple graphical method was used to identify the most efficient fluxes for conversion of substrate into energy and biomass for all culturing conditions between oxygen sufficiency and the complete absence of oxygen.

The elementary mode analysis method provides a very detailed description of the robustness of the *E. coli* network.

For the different culturing conditions, it was possible to view all of the possible biomass producing pathways. For oxygen-sufficient conditions, the 200-min doubling time culture had 19 unique pathways that produced biomass with carbon yields ranging from 0.80 and 0.82. By comparing these pathways it was possible to identify which reactions were necessary for high yielding aerobic growth. While this network robustness permits the *E. coli* network to effectively handle different culturing stresses and a multitude of substrates, the large number of biomass-producing pathways can complicate the analysis, control, and engineering of experimental systems. For example, the potential benefits of engineering a target enzyme or pathway may not be realized if a large metabolite flux uses an alternative pathway that by-passes the modified site(s). The presented analysis technique could be used to identify knockout targets that would require flux through the modified pathway(s). This system would likely be more responsive to genetic modifications like gene overexpression since metabolite flux through the modified enzyme would be required for the culture to grow. In addition, from a control standpoint, having only one possible flux state should reduce experimental variations in parameters like yields and metabolite flux patterns.

Observed glucose consumption rates can be described as the sum of two contributions. One is specified by the stoichiometry of the operational biomass elementary mode and the growth rate. The second contribution accounts for maintenance energy and its magnitude has to be derived from experimental data. However, glucose consumption rates correlate linearly with growth rate (Schultz and Lipe, 1964). Therefore, the rate of glucose consumed by the energy mode can be estimated. This is important because recently methods have been developed that can probe the rate of glucose uptake at the individual cell level (Natarajan and Srienc, 1999). Thus, the elementary-mode analysis can be extended to the single-cell level to estimate rates of each individual reaction occurring in each individual cell of a population.

Identifying the most efficient pathways for energy and biomass production defines the simplest flux unit available to cells under steady-state conditions. These units are powerful tools that can be used to analyze, interpret, and predict the behavior of *E. coli* under a wide variety of culturing conditions. For instance, it permits the evaluation of the entire internal rate structure of *E. coli*. This will be presented in a forthcoming communication (Carlson and Srienc, submitted).

Ross Carlson is a recipient of a NIH training grant fellowship in Biotechnology.

NOMENCLATURE

Metabolite abbreviations

ACETATE	acetate
ACETATE_ext	extracellular acetate
ACETYL_CoA	acetyl-coenzyme A
ADP	adenosine diphosphate
AKG	alpha-ketoglutarate
ATP	adenosine triphosphate
ATP_main	maintenance energy
BIOMASS	biomass
CITRATE	citrate
CO2	carbon dioxide
CO2_ext	extracellular carbon dioxide
CoASH	coenzyme A
DHAP	dihydroxyacetone phosphate
ERYTH_4_P	erythrose-4-phosphate
ETOH	ethanol
ETOH_ext	extracellular ethanol
FAD	flavin adenine dinucleotide
FADH	flavin adenine dinucleotide
FORMATE	formate
FORMATE_ext	extracellular formate
FRU_6_P	fructose-6-phosphate
FUMARATE	fumarate
GA_3P	glyceraldehyde-3-phosphate
GLU_6_P	glucose-6-phosphate
GLC_ext	extracellular glucose
ISOCIT	isocitrate
LACTATE	lactate
LACTATE_ext	extracellular lactate
MALATE	malate
NAD	nicotinamide adenine dinucleotide
NADH	nicotinamide adenine dinucleotide
NH3	ammonia
NH3_ext	extracellular ammonia
OXALO	oxaloacetate
OXY_ext	extracellular monooxygen
PEP	phosphoenolpyruvate
PG	phosphoglycerate
PYR	pyruvate
RIBOSE_5_	ribose-5-phosphate
RIBULOSE_5_P	ribulose-5-phosphate
SED_7_P	sedoheptulose-7-P
SUCC	succinate
SUCC_CoA	succinyl-coenzyme A
SUCC_ext	extracellular succinate
XYL_5_P	xylulose-5-phosphate

APPENDIX A.1

Escherichia coli metabolic network model input file used for the program METATOOL. The 'Reactions' section lists all considered reactions and their stoichiometries.

The metabolite coefficients for biomass generation (reaction R70) are listed in Table II.

-ENZREV (reversible reactions)
R2r R5r R6r R7r R8r R11r R12r R13r R14r R15r R23r R26r R27r R28r R29r R53r R54r R97r

-ENZIRREV (irreversible reactions)
R1 R3 R4 R9 RR9 R10 R20 R21 R22 R24 R25 R40 R41 R42 R55 R70 R80 R81 R82 R83 R90 R91 R93 R94 R95 R96

-METINT (internal metabolite declaration)
ATP ADP GLU_6_P FRU_6_P FRU_BIS_P DHAP GA_3P NAD NADH

RIBULOSE_5_P XYL_5_P RIBOSE_5_P SED_7_P ERYTH_4_P PYR PEP CITRATE
OXALO MALATE CoASH ACETYL_CoA FADH FAD AKG ISOCIT ACETATE SUCC FUMARATE PG LACTATE SUCC_CoA NH3 ETOH FORMATE CO2 GLOXY

-METEXT (external metabolite declaration)
ATP_main GLU_ext ETOH_ext ACETATE_ext CO2_ext LACTATE_ext SUCC_ext NH3_ext FORMATE_ext BIOMASS OXY_ext

Reactions:

-Glycolysis
R1: GLU_ext + PEP = GLU_6_P + PYR.
R2r: GLU_6_P = FRU_6_P.
R3: FRU_6_P + ATP = FRU_BIS_P + ADP.
R4: FRU_BIS_P = FRU_6_P.
R5r: FRU_BIS_P = DHAP + GA_3P.
R6r: GA_3P = DHAP.
R7r: GA_3P + ADP + NAD = PG + ATP + NADH.
R8r: PG = PEP.
R9: PEP + ADP = PYR + ATP.
RR9: PYR + 2 ATP = PEP + 2 ADP.

-Pentose Phosphate Pathway
R10: GLU_6_P + 2 NAD = RIBULOSE_5_P + 2 NADH + CO2.
R11r: RIBULOSE_5_P = XYL_5_P.
R12r: RIBULOSE_5_P = RIBOSE_5_P.
R13r: RIBOSE_5_P + XYL_5_P = SED_7_P + GA_3P.
R14r: GA_3P + SED_7_P = ERYTH_4_P + FRU_6_P.
R15r: ERYTH_4_P + XYL_5_P = GA_3P + FRU_6_P.

-TCA cycle
R20: PYR + CoASH = ACETYL_CoA + FORMATE.
R21: PYR + NAD + CoASH = ACETYL_CoA + CO2 + NADH.
R22: OXALO + ACETYL_CoA = CITRATE + CoASH.
R23r: CITRATE = ISOCIT.
R24: ISOCIT + NAD = AKG + NADH + CO2.
R25: AKG + NAD + CoASH = NADH + SUCC_CoA + CO2.
R26r: SUCC_CoA + ADP = SUCC + ATP + CoASH.
R27r: SUCC + FAD = FUMARATE + FADH.
R28r: FUMARATE = MALATE.
R29r: MALATE + NAD = OXALO + NADH.

-Anapleurotic reactions
R40: PEP + CO2 = OXALO.
R41: MALATE + NAD = PYR + NADH + CO2.
R42: OXALO + ATP = PEP + ADP + CO2.

-Redox-associated reactions
R53r: PYR + NADH = LACTATE + NAD.
R54r: ACETYL_CoA + 2 NADH = ETOH + 2 NAD + CoASH.
R55: ACETYL_CoA + ADP = ACETATE + CoASH + ATP.

-Biomass production (see Table II for doubling time dependent reaction coefficients)
R70: * GLU_6_P + * RIBOSE_5_P + * ERYTH_4_P + * PEP + * PYR + * ACETYL_CoA + * AKG + * OXALO + * ATP + * NADH + * NH3 = *BIOMASS + * CoASH + * ADP + * NAD + *CO2.

-Oxidative phosphorylation/maintenance energy:
R80: NADH + 2 ADP + OXY_ext = NAD + 2 ATP.
R81: FADH + ADP + OXY_ext = FAD + ATP.
R82: ATP = ADP + ATP_main.
R83: NADH + FAD = NAD + FADH.

-Membrane transport reactions
R90: ETOH = ETOH_ext.
R91: ACETATE = ACETATE_ext.
R93: NH3_ext = NH3.
R94: LACTATE = LACTATE_ext.
R95: SUCC = SUCC_ext.
R96: FORMATE = FORMATE_ext.
R97r: CO2 = CO2_ext.

APPENDIX A.2

Relative flux patterns for ATP-producing modes. Each individual reaction involved in the mode is listed. The number preceding the reaction number is the relative flux through that reaction. When there is no preceding number, the flux is unity. The reactions are in the format of the METATOOL output file, *i.e.* (2 R26r) designates reaction 26 which is reversible and has a relative flux of 2. See Appendix A.1 for the stoichiometry of each reaction. See Table III for the overall stoichiometry of the external metabolites of each mode.

Mode:	Reaction set:
M_1^{ATP}	R1 R2r R3 R5r -R6r (2 R7r) (2 R8) R9 (2 R21) (2 R22) (2 R23r) (2 R24) (2 R25) (2 R26r) (2 R27r) (2 R28r) (2 R29r) (10 R80) (2 R81) (26 R83) (6 R97r)
M_2^{ATP}	R1 R2r R3 R5r -R6r (2 R7r) (2 R8) R9 (2 R21) (2 R55) (4 R80) (12 R83) (2 R91) (2 R97r)
M_3^{ATP}	R1 R2r R3 R5r -R6r (2 R7r) (2 R8) R9 (2 R20) (2 R55) (2 R80) (8 R83) (2 R91) (2 R96)
M_4^{ATP}	R1 R2r R3 R5r -R6r (2 R7r) (2 R8) R9 (2 R20) R54r R55 (3 R83) R90 R91 (2 R96)

APPENDIX A.3

Carbon and oxygen yields and overall stoichiometry for the biomass producing modes representing metabolic states 1–4 and cultures with doubling times of 200, 100, 80, 60, 50, 40, and 30 minutes. See Appendix A.4 for the relative flux patterns for each mode. Carbon yield is defined as the ratio of carbon atoms in the biomass by the number of substrate (glucose) carbon atoms. Oxygen yield is defined as the ratio of carbon atoms in biomass by the number of O_2 moles required to produce the biomass. The biomass terms listed below can be converted into C moles of biomass by multiplying them by the number of carbon atoms per biomass term listed in Table II.

Mode:	Doubling time (min)	$Y_{x/C}$	$Y_{x/O2}$	Mode stoichiometry:
$M_1^{\times 200}$	200	0.821	4.36	14005 glucose + 19006 NH_3 + 15806 O_2 = 15078 CO_2 + 26 biomass

Mode:	Doubling time (min)	$Y_{x/C}$	$Y_{x/O2}$	Mode stoichiometry:
$M_2^{\times 200}$	200	0.689	5.74	7691 glucose + 8772 NH_3 + 5542 O_2 = 4558 acetate + 5206 CO_2 + 12 biomass
$M_3^{\times 200}$	200	0.570	8.10	6203 glucose + 832 CO_2 + 5848 NH_3 + 2619 O_2 = 5190 acetate + 6454 formate + 8 biomass
$M_4^{\times 200}$	200	0.335	anaerobic	1321 glucose + 104 CO_2 + 731 NH_3 = 873 ethanol + 867 acetate + 1898 formate + biomass
$M_1^{\times 100}$	100	0.823	4.45	8181 glucose + 11024 NH_3 + 9085 O_2 = 8682 CO_2 + 26 biomass
$M_2^{\times 100}$	100	0.696	5.81	372 glucose + 424 NH_3 + 267.5 O_2 = 213 acetate + 252 CO_2 + biomass
$M_3^{\times 100}$	100	0.575	8.23	1803 glucose + 252 CO_2 + 1696 NH_3 + 755 O_2 = 1482 acetate + 1890 formate + 4 biomass
$M_4^{\times 100}$	100	0.341	anaerobic	2296 glucose + 189 CO_2 + 1272 NH_3 = 1510 ethanol + 1489 acetate + 3305 formate + 3 biomass
$M_1^{\times 80}$	80	0.822	4.56	17245 glucose + 22698 NH_3 + 18645 O_2 = 18372 CO_2 + 78 biomass
$M_2^{\times 80}$	80	0.701	5.92	1557 glucose + 1746 NH_3 + 1105 O_2 = 856 acetate + 1084 CO_2 + 6 biomass
$M_3^{\times 80}$	80	0.577	8.51	3785 glucose + 516 CO_2 + 3492 NH_3 + 1539 O_2 = 3054 acetate + 4026 formate + 12 biomass
$M_4^{\times 80}$	80	0.344	anaerobic	3175 glucose + 258 CO_2 + 1746 NH_3 = 2052 ethanol + 2040 acetate + 4578 formate + 6 biomass
$M_1^{\times 60}$	60	0.825	4.81	2040 glucose + 2574 NH_3 + 2100 O_2 = 2139 CO_2 + 13 biomass
$M_2^{\times 60}$	60	0.712	6.18	2183 glucose + 2376 NH_3 + 1510 O_2 = 1114 acetate + 1546 CO_2 + 12 biomass
$M_3^{\times 60}$	60	0.584	9.02	1773 glucose + 240 CO_2 + 1584 NH_3 + 689 O_2 = 1378 acetate + 1906 formate + 8 biomass
$M_4^{\times 60}$	60	0.355	anaerobic	2191 glucose + 180 CO_2 + 1188 NH_3 = 1378 ethanol + 1378 acetate + 3152 formate + 6 biomass
$M_1^{\times 50}$	50	0.813	4.86	10759 glucose + 13338 NH_3 + 10812 O_2 = 11046 CO_2 + 78 biomass
$M_2^{\times 50}$	50	0.707	6.17	476 glucose + 513 NH_3 + 327 O_2 = 231 acetate + 336 CO_2 + 3 biomass
$M_3^{\times 50}$	50	0.580	9.06	2321 glucose + 324 CO_2 + 2052 NH_3 + 891 O_2 = 1758 acetate + 2502 formate + 12 biomass
$M_4^{\times 50}$	50	0.354	anaerobic	1903 glucose + 162 CO_2 + 1026 NH_3 = 1188 ethanol + 1176 acetate + 2736 formate + 6 biomass
$M_1^{\times 40}$	40	0.687	4.18	9421 glucose + 11544 NH_3 + 9297 O_2 = 9804 CO_2 + 78 biomass
$M_2^{\times 40}$	40	0.600	5.30	1661 glucose + 1776 NH_3 + 1128 O_2 = 786 acetate + 1206 CO_2 + 12 biomass
$M_3^{\times 40}$	40	0.490	7.89	4063 glucose + 552 CO_2 + 3552 NH_3 + 1515 O_2 = 3054 acetate + 4446 formate + 24 biomass
$M_4^{\times 40}$	40	0.302	anaerobic	1647 glucose + 138 CO_2 + 888 NH_3 = 1010 ethanol + 1016 acetate + 2374 formate + 6 biomass
$M_1^{\times 30}$	30	0.690	4.21	4439 glucose + 5421 NH_3 + 4365 O_2 = 4599 CO_2 + 39 biomass
$M_2^{\times 30}$	30	0.602	5.33	391 glucose + 417 NH_3 + 265 O_2 = 184 acetate + 283 CO_2 + 3 biomass
$M_3^{\times 30}$	30	0.492	7.95	1913 glucose + 264 CO_2 + 1668 NH_3 + 711 O_2 = 1434 acetate + 2094 formate + 12 biomass
$M_4^{\times 30}$	30	0.304	anaerobic	1549 glucose + 132 CO_2 + 834 NH_3 = 948 ethanol + 954 acetate + 2232 formate + 6 biomass

APPENDIX A.4

Relative flux patterns for biomass producing modes representing *E. coli* cultures with doubling times of 200, 100, 80, 60, 50, 40, and 30 minutes. See Appendix A.1 for the designation of each reaction. See Appendix A.3 for the overall stoichiometry of each mode. The superscript refers to the specific growth rate given as a doubling time in minutes, i.e., $M_1^{\times 200}$ designates a 200-min doubling time. A schematic representation of the carbon fluxes is shown in Figure 4.

Mode:	Reaction set:
$M_1^{\times 200}$	(14005 R1) (13901 R2r) (12835 R3) (12835 R5r) (-12835 R6r) (24734 R7r) (24734 R8r) (823 R9r) (-1066 R11r) (1066 R12r) (-130 R13r) (-130 R14r) (-936 R15r) (8666 R21) (6794 R22) (6794 R23r) (6794 R24) (4558 R25) (4558 R26r) (4558 R27r) (4558 R28r) (4558 R29r) (5850 R40) (26 R70) (27054 R80) (4558 R81) (19006 R93) (15078 R97r)
$M_2^{\times 200}$	(7691 R1) (7643 R2r) (7151 R3) (7151 R5r) (-7151 R6r) (13870 R7r) (13870 R8r) (1607 R9r) (-492 R11r) (492 R12r) (-60 R13r) (-60 R14r) (-432 R15r) (6454 R21) (1032 R22) (1032 R23r) (1032 R24) (2700 R40) (4558 R55) (12 R70) (11084 R80) (4558 R91) (8772 R93) (5206 R97r)
$M_3^{\times 200}$	(6203 R1) (6171 R2r) (5843 R3) (5843 R5r) (-5843 R6r) (11398 R7r) (11398 R8r) (2147 R9r) (-328 R11r) (328 R12r) (-40 R13r) (-40 R14r) (-288 R15r) (6454 R20) (688 R22) (688 R23r) (688 R24) (1800 R40) (5190 R55) (8 R70) (5238 R80) (5190 R91) (5848 R93) (6454 R96 (-832 R97r))

Mode:	Reaction set:
$M_4^{\times 200}$	(1321 R1) (1317 R2r) (1276 R3) (1276 R5r) (-1276 R6r) (2516 R7r) (2516 R8r) (814 R9r) (-41 R11r) (41 R12r) (-5 R13r) (-5 R14r) (-36 R15r) (1898 R20) (86 R22) (86 R23r) (86 R24) (225 R40) (873 R54r) (867 R55) R70 (873 R90) (867 R91) (731 R93) (1898 R96) (-104 R97r)
$M_1^{\times 100}$	(8181 R1) (8077 R2r) (7375 R3) (7375 R5r) (-7375 R6r) (14178 R7r) (14178 R8r) (381 R9r) (-702 R11r) (702 R12r) (-130 R13r) (-130 R14r) (-572 R15r) (5208 R21) (3778 R22) (3778 R23r) (3778 R24) (2556 R25) (2556 R26r) (2556 R27r) (2556 R28r) (2556 R29r) (3250 R40) (26 R70) (15614 R80) (2556 R81) (11024 R93) (8682 R97r)
$M_2^{\times 100}$	(372 R1) (368 R2r) (341 R3) (341 R5r) (-341 R6r) (660 R7r) (660 R8r) (72 R9r) (-27 R11r) (27 R12r) (-5 R13r) (-5 R14r) (-22 R15r) (315 R21) (47 R22) (47 R23r) (47 R24) (125 R40) (213 R55) R70 (535 R80) (213 R91) (424 R93) (252 R97r)
$M_3^{\times 100}$	(1803 R1) (1787 R2r) (1679 R3) (1679 R5r) (-1679 R6r) (3270 R7r) (3270 R8r) (603 R9r) (-108 R11r) (108 R12r) (-20 R13r) (-20 R14r) (-88 R15r) (1890 R20) (188 R22) (188 R23r) (188 R24) (500 R40) (1482 R55) (4 R70) (1510 R80) (1482 R91) (1696 R93) (1890 R96) (-252 R97r)
$M_4^{\times 100}$	(2296 R1) (2284 R2r) (2203 R3) (2203 R5r) (-2203 R6r) (4340 R7r) (4340 R8r) (1396 R9r) (-81 R11r) (81 R12r) (-15 R13r) (-15 R14r) (-66 R15r) (3305 R20) (141 R22) (141 R23r) (141 R24) (375 R40) (1510 R54r) (1489 R55) (3 R70) (150 R90) (1489 R91) (1272 R93) (3305 R96) (-189 R97r)
$M_1^{\times 80}$	(17245 R1) (16933 R2r) (15503 R3) (15503 R5r) (-15503 R6r) (29862 R7r) (29862 R8r) (995 R9r) (-1430 R11r) (143 R12r) (-286 R13r) (-286 R14r) (-1144 R15r) (11454 R21) (7632 R22) (7623 R23r) (7632 R24) (5136 R25) (5136 R26r) (5136 R27r) (5136 R28r) (5136 R29r) (6630 R40) (78 R70) (32154 R80) (5136 R81) (22698 R93) (18372 R97r)
$M_2^{\times 80}$	(1557 R1) (1533 R2r) (1423 R3) (1423 R5r) (-1423 R6r) (2758 R7r) (2758 R8r) (307 R9r) (-110 R11r) (110 R12r) (-22 R13r) (-22 R14r) (-88 R15r) (1342 R21) (192 R22) (192 R23r) (192 R24) (510 R40) (856 R55) (6 R70) (2210 R80) (856 R91) (1746 R93) (1084 R97r)
$M_3^{\times 80}$	(3785 R1) (3737 R2r) (3517 R3) (3517 R5r) (-3517 R6r) (6858 R7r) (6858 R8r) (1285 R9r) (-220 R11r) (220 R12r) (-44 R13r) (-44 R14r) (-176 R15r) (4026 R20) (384 R22) (384 R23r) (384 R24) (1020 R40) (3054 R55) (12 R70) (3078 R80) (3054 R91) (3492 R93) (4026 R96) (-516 R97r)
$M_4^{\times 80}$	(3175 R1) (3151 R2r) (3041 R3) (3041 R5r) (-3041 R6r) (5994 R7r) (5994 R8r) (1925 R9r) (-110 R11r) (110 R12r) (-22 R13r) (-22 R14r) (-88 R15r) (4578 R20) (92 R22) (192 R23r) (192 R24) (510 R40) (2052 R54r) (2040 R55) (6 R70) (2052 R90) (2040 R91) (1746 R93) (4578 R96) (-258 R97r)
$M_1^{\times 60}$	(2040 R1) (1988 R2r) (1819 R3) (1819 R5r) (-1819 R6r) (3508 R7r) (3508 R8r) (129 R9r) (-169 R11r) (169 R12r) (-39 R13r) (-39 R14r) (-130 R15r) (1415 R21) (843 R22) (843 R23r) (843 R24) (557 R25) (557 R26r) (557 R27r) (557 R28r) (557 R29r) (754 R40) (13 R70) (3643 R80) (557 R81) (2574 R93) (2139 R97r)
$M_2^{\times 60}$	(2183 R1) (2135 R2r) (1979 R3) (1979 R5r) (-1979 R6r) (3838 R7r) (3838 R8r) (419 R9r) (-156 R11r) (156 R12r) (-36 R13r) (-36 R14r) (-120 R15r) (1906 R21) (264 R22) (264 R23r) (264 R24) (696 R40) (1114 R55) (12 R70) (3020 R80) (1114 R91) (2376 R93) (1546 R97r)
$M_3^{\times 60}$	(1773 R1) (1741 R2r) (1637 R3) (1637 R5r) (-1637 R6r) (3194 R7r) (3194 R8r) (597 R9r) (-104 R11r) (104 R12r) (-24 R13r) (-24 R14r) (-80 R15r) (1906 R20) (176 R22) (176 R23r) (176 R24) (464 R40) (1378 R55) (8 R70) (1378 R80) (1378 R91) (1584 R93) (1906 R96) (-240 R97r)
$M_4^{\times 60}$	(2191 R1) (2167 R2r) (2089 R3) (2089 R5r) (-2089 R6r) (4118 R7r) (4118 R8r) (1309 R9r) (-78 R11r) (78 R12r) (-18 R13r) (-18 R14r) (-60 R15r) (3152 R20) (132 R22) (132 R23r) (132 R24) (348 R40) (1378 R54r) (1378 R55) (6 R70) (1378 R90) (1378 R91) (1188 R93) (3152 R96) (-180 R97r)
$M_1^{\times 50}$	(10759 R1) (10447 R2r) (9563 R3) (9563 R5r) (-9563 R6r) (18450 R7r) (18450 R8r) (749 R9r) (-884 R11r) (884 R12r) (-208 R13r) (-208 R14r) (-676 R15r) (7608 R21) (4254 R22) (4254 R23r) (4254 R24) (2772 R25) (2772 R26r) (2772 R27r) (2772 R28r) (2772 R29r) (3900 R40) (78 R70) (18852 R80) (2772 R81) (13338 R93) (11046 R97r)
$M_2^{\times 50}$	(476 R1) (464 R2r) (430 R3) (430 R5r) (-430 R6r) (834 R7r) (834 R8r) (91 R9r) (-34 R11r) (34 R12r) (-8 R13r) (-8 R14r) (-26 R15r) (417 R21) (57 R22) (27 R23r) (57 R24) (150 R40) (231 R55) (3 R70) (654 R80) (231 R91) (513 R93) (336 R97r)
$M_3^{\times 50}$	(2321 R1) (2273 R2r) (2137 R3) (2137 R5r) (-2137 R6r) (4170 R7r) (4170 R8r) (781 R9r) (-136 R11r) (136 R12r) (-32 R13r) (-32 R14r) (-104 R15r) (2502 R20) (228 R22) (228 R23r) (228 R24) (600 R40) (1758 R55) (12 R70) (1782 R80) (1758 R91) (2052 R93) (2502 R96) (-324 R97r)
$M_4^{\times 50}$	(1903 R1) (1879 R2r) (1811 R3) (1811 R5r) (-1811 R6r) (3570 R7r) (3570 R8r) (133 R9r) (-68 R11r) (68 R12r) (-16 R13r) (-16 R14r) (-52 R15r) (2736 R20) (114 R22) (114 R23r) (114 R24) (300 R40) (1188 R54r) (1176 R55) (6 R70) (1188 R90) (1176 R91) (1026 R93) (2736 R96) (-162 R97r)
$M_1^{\times 40}$	(9421 R1) (9109 R2r) (8303 R3) (8303 R5r) (-8303 R6r) (16008 R7r) (16008 R8r) (659 R9r) (-806 R11r) (806 R12r) (-208 R13r) (-208 R14r) (-598 R15r) (6882 R21) (3606 R22) (3606 R23r) (3606 R24) (2358 R25) (2358 R26r) (2358 R27r) (2358 R28r) (2358 R29r) (3276 R40) (78 R70) (16236 R80) (2358 R81) (11544 R93) (9804 R97r)
$M_2^{\times 40}$	(1661 R1) (1613 R2r) (1489 R3) (1489 R5r) (-1489 R6r) (2886 R7r) (2886 R8r) (313 R9r) (-124 R11r) (124 R12r) (-32 R13r) (-32 R14r) (-92 R15r) (1482 R21) (192 R22) (192 R23r) (192 R24) (504 R40) (786 R55) (12 R70) (2256 R80) (786 R91) (1776 R93) (1206 R97r)
$M_3^{\times 40}$	(4063 R1) (3967 R2r) (3719 R3) (3719 R5r) (-3719 R6r) (7254 R7r) (7254 R8r) (1367 R9r) (-248 R11r) (248 R12r) (-64 R13r) (-64 R14r) (-184 R15r) (4446 R20) (384 R22) (384 R23r) (384 R24) (1008 R40) (3054 R55) (24 R70) (3030 R80) (3054 R91) (3552 R93) (4446 R96) (-552 R97r)
$M_4^{\times 40}$	(1647 R1) (1623 R2r) (1561 R3) (1561 R5r) (-1561 R6r) (3076 R7r) (3076 R8r) (973 R9r) (-62 R11r) (62 R12r) (-16 R13r) (-16 R14r) (-46 R15r) (2374 R20) (96 R22) (96 R23r) (96 R24) (252 R40) (1010 R54r) (1016 R55) (6 R70) (1010 R90) (1016 R91) (888 R93) (2374 R96) (-138 R97r)
$M_1^{\times 30}$	(4439 R1) (4283 R2r) (3880 R3) (3880 R5r) (-3880 R6r) (7461 R7r) (7461 R8r) (292 R9r) (-403 R11r) (403 R12r) (-104 R13r) (-104 R14r) (-299 R15r) (3249 R21) (1650 R23r) (1650 R24) (1104 R25) (1104 R26r) (1104 R27r) (1104 R28r) (1104 R29r) (1650 R22) (1482 R40) (39 R70) (7626 R80) (1104 R81) (5421 R93) (4599 R97r)
$M_2^{\times 30}$	(391 R1) (379 R2r) (348 R3) (348 R5r) (-348 R6r) (673 R7r) (673 R8r) (72 R9r) (-31 R11r) (31 R12r) (-8 R13r) (-8 R14r) (-23 R15r) (349 R21) (42 R22) (42 R23r) (42 R24) (114 R40) (184 R55) (3 R70) (530 R80) (184 R91) (417 R93) (283 R97r)
$M_3^{\times 30}$	(1913 R1) (1865 R2r) (1741 R3) (1741 R5r) (-1741 R6r) (3390 R7r) (3390 R8r) (637 R9r) (-124 R11r) (124 R12r) (-32 R13r) (-32 R14r) (-92 R15r) (2094 R20) (168 R22) (168 R23r) (168 R24) (456 R40) (1434 R55) (12 R70) (1422 R80) (1434 R91) (1668 R93) (2094 R96) (-264 R97r)
$M_4^{\times 30}$	(1549 R1) (1525 R2r) (1463 R3) (1463 R5r) (-1463 R6r) (2880 R7r) (2880 R8r) (911 R9r) (-62 R11r) (62 R12r) (-16 R13r) (-16 R14r) (-46 R15r) (2232 R20) (84 R22) (84 R23r) (84 R24) (228 R40) (948 R54r) (954 R55) (6 R70) (948 R90) (954 R91) (834 R93) (2232 R96) (-132 R97r)

References

Alam KY, Clark DP. 1989. Anaerobic fermentation balance of *Escherichia coli* as observed by in vivo nuclear magnetic resonance spectroscopy. J Bacteriol 171(11):6213–6217.

Alexeeva S, de Kort B, Sawers G, Hellingwerf KJ, Teixeira de Mattos MJ. 2000. Effects of limited aeration and of the rcAB system on intermediary pyruvate catabolism in *Escherichia coli*. J Bacteriol 182(17):4934–4940.

Amarasingham CR, Davis BD. 1965. Regulation of α-ketoglutarate dehydrogenase formation in *Escherichia coli*. J Biol Chem 240(9):3664–3668.

Belaich A, Belaich JP. 1976. Microcalorimetric study of the anaerobic growth of *Escherichia coli*: Growth thermograms in a synthetic medium. J Bacteriol 125(1):14–18.

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.

Boonstra B, French CE, Wainwright I, Bruce NC. 1999. The *udhA* gene of *Escherichia coli* encodes a soluble pyridine nucleotide transhydrogenase. J Bacteriol 181(3):1030–1034.

Bragg PD, Davies PL, Hou C. 1972. Function of energy-dependent transhydrogenase in *Escherichia coli*. Biochem Biophys Res Commun 47(5):1248–1255.

Bremer H, Dennis PP. 1987. Modulation of chemical composition and other parameters of the cell by growth rate. In: Neidhardt FC, editor. *Escherichia coli* and *Salmonella typhimurium:* Cellular and molecular biology. Washington D.C.: American Society for Microbiology. p 1527–1542.

Burgard AP, Maranas CD. 2001. Probing the performance limits of the *Escherichia coli* metabolic network subject to gene additions or deletions. Biotechnol Bioeng 74(5):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(10): 3025–3029.

Carlson R, Fell D, Srienc F. 2002. Metabolic pathway analysis of a recombinant yeast for rational strain development. Biotechnol Bioeng 79(2):121–134.

Chesbro W, Evans T, Eifert R. 1979. Very slow growth of *Escherichia coli*. J Bacteriol 139(2):625–638.

Churchward G, Bremer H, Young R. 1982. Macromolecular composition of bacteria. J Theor Biol 94:651–670.

Conradt H, Hohmann-Berger M, Hohman H-P, Blaschkowski HP, Knappe J. 1983. Pyruvate formate-lyase (inactive form) and pyruvate formate-lyase activating enzyme of *Escherichia coli*: Isolation and structural properties. Arch Biochem Biophys 228(1): 133–142.

Coomes MW, Mitchell BK, Beezley A, Smith TE. 1985. Properties of an *Escherichia coli* mutant deficient in phosphoenolpyruvate carboxylase catalytic activity. J Bacteriol 164(2):646–652.

Dandekar T, Schuster S, Snel B, Huynen M, Bork P. 1999. Pathway alignment: Application to the comparative analysis of glycolytic enzymes. Biochem J 343(Part 1):115–124.

Dennis PP, Bremer H. 1974. Macromolecular composition during steady-state growth of *Escherichia coli*. Br J Bacteriol 119(1):270–281.

Edwards JS, Palsson BO. 2000a. Metabolic flux balance analysis and the in silico analysis of *Escherichia coli* K-12 gene deletions. BMC Bioinformatics 1(1):1.

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

Edwards JS, Ramakrishna R, Palsson BO. 2002. Characterizing the metabolic phenotype: A phenotype phase plane analysis. Biotechnol Bioeng 77(1):27–36.

Farmer IS, Jones CW. 1976. The effect of temperature on the molar growth yield and maintenance. FEBS Lett. 67(3):359–363.

Fiaux J, Andersson CIJ, Holmberg N, Bulow L, Kallio P, Szyperski T, Bailey JE, Wuthrich K. 1999. ^{13}C NMR flux ration analysis of *Escherichia coli* central carbon metabolism in microaerobic bioprocesses. J Am Chem Soc 121:1407–1408.

Fraenkel DG. 1968. Selection of *Escherichia coli* mutants lacking glucose-6-phosphate dehydrogenase or gluconate-6-phosphate dehydrogenase. J Bacteriol 95(4):1267–1271.

Gennis RB, Stewart V. 1996. Respiration. In: Neidhardt FC, editor. *Escherichia coli* and *Salmonella typhimurium*. Washington DC: American Society for Microbiology. p 217–261.

Hansen HG, Henning U. 1966. Regulation of pyruvate dehydrogenase activity in *Escherichia coli* K12. Biochim Biophys Acta 122(2): 355–358.

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(3):1076–1087.

Hirsch CA, Rasminsky M, Davis BD, Lin EC. 1963. A fumarate reductase in *Escherichia coli* distinct from succinate dehydrogenase. J Biol Chem 238(11):3770–3774.

Holms WH, Bennett PM. 1971. Regulation of isocitrate dehydrogenase activity in *Escherichia coli* on adaptation to acetate. J Gen Microbiol 65(1):57–68.

Ingraham JL, Maaløe O, Neidhardt FC. 1983. Growth of the bacterial cell. Sunderland, MA: Sinauer Associates, Inc. p 1–172.

Klamt S, Stelling J. 2003. Two approaches for metabolic pathway analysis? Trends Biotechnol 21(2): 64–69.

Knappe J. 1987. Anaerobic dissimilation of pyruvate. In: Neidhardt FC, editor. *Escherichia coli* and *Salmonella typhimurium*. Washington DC: American Society for Microbiology. p 151–155.

Koch AL. 1971. The adaptive responses of *Escherichia coli* to a feast and famine existence. Adv Microb Physiol 6:147–217.

Kornberg HL. 1966a. The role and control of the glyoxylate cycle in *Escherichia coli*. Biochem J 99(1):1–11.

Kornberg HL. 1966b. The role and maintenance of the tricarboxylic acid cycle in *Escherichia coli*. Biochem Soc Symp 30:155–171.

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. New York: Marcel Dekker Inc.

Liao JC, Hou S-Y, Chao Y-P. 1996. Pathway analysis, engineering and physiological considerations for redirecting central metabolism. Biotechnol Bioeng 52:129–140.

Lin ECC, Kuritzkes DR. 1987. Pathways for anaerobic electron transport. In: Neidhardt FC, editor. *Escherichia coli* and *Salmonella typhimurium*. Washington DC: American Society for Microbiology. p 201–221.

Majewski RA, Domach MM. 1990. Simple constrained-optimization view of acetate overflow in *E. coli*. Biotechnol Bioeng 35:732–738.

Natarajan A, Srienc F. 1999. Dynamics of glucose uptake by single *Escherichia coli* cells. Metab Eng 1(4):320–333.

Neidhardt FC, editor. 1987. *Escherichia coli* and *Salmonella typhimurium*. Washington DC: American Society for Microbiology.

Neidhardt FC, editor. 1996. *Escherichia coli* and *Salmonella*, 2nd ed. Washington DC: American Society for Microbiology.

Neijssel OM, Tempest DW. 1976. Bioenergetic aspects of aerobic growth of *Klebsiella aerogenes* NCTC 418 in carbon-limited and carbon-sufficient chemostat culture. Arch Microbiol 107(2):215–221.

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.

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(1):67–82.

Pfeiffer T, Sanchez-Valdenebro I, Nuno JC, Montero F, Schuster S. 1999. METATOOL: For studying metabolic networks. Bioinformatics 15(3):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(2):124–37.

Pirt SJ. 1965. The maintenance energy of bacteria in growing cultures. Proc R Soc Lond B Biol Sci 163(991):224–231.

Pramanik J, Keasling JD. 1997. Effect of *Escherichia coli* biomass composition on central metabolic fluxes predicted by a stoichiometric model. Biotechnol Bioeng 60(2):230–238.

Price ND, Papin JA, Palsson B. 2002. Determination of redundancy and systems properties of the metabolic network of *Helicobacter pylori* using genome scale extreme pathway analysis. Genome 12(5): 760–769.

Schilling CH, Edwards JS, Letscher D, Palsson BO. 2000a. Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems. Biotechnol Bioeng 71(4):286–306.

Schilling CH, Schuster S, Palsson BO, Heinrich R. 1999. Metabolic pathway analysis: Basic concepts and scientific applications in the post-genomic 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 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.

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(3):326–332.

Schuster S, Hilgetag C, Fell D. 1994. Detecting elementary modes of functioning in metabolic networks. Mod 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(2):153–181.

Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED. 2002. Metabolic network structure determines key aspects of functionality and regulation. Nature 420(6912):190–193.

Stouthamer AH, Bettenhaussen C. 1973. Utilization of energy for growth and maintenance in continuous and batch cultures of microorganisms. A reevaluation of the method for the determination of ATP production by measuring molar growth yields. Biochim Biophys Acta 301(1):53–70.

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.

Thomas AD, Doelle HW, Westwood AW, Gordon GL. 1972. Effect of oxygen on several enzymes involved in the aerobic and anaerobic utilization of glucose in *Escherichia coli*. J Bacteriol 112(3): 1099–1105.

Unden G, Bongaerts J. 1997. Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. Biochim Biophys Acta 1320:217–234.

Vallino J, Stephanopoulos G. 1990. Flux determinations in cellular bioreaction networks: Applications to lysine fermentations. In: Sikdar SK, Bier M, Todd P, editors. Frontiers in bioprocessing. Boca Raton, FL: CRC Press. p 205–219.

van de Walle M, Shiloach J. 1998. Proposed mechanism of acetate accumulation in two recombinant *Escherichia coli* strains during high density fermentation. Biotechnol Bioeng 57(1):71–78.

Van Dien SJ, Lidstrom ME. 2002. Stoichiometric model for evaluating the metabolic capabilities of the facultative methylotroph *Methylobacterium extorquens* AM1, with application to reconstruction of C(3) and C(4) metabolism. Biotechnol Bioeng 78(3):296–312.

Varma A, Boesch BW, Palsson BO. 1993. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. Appl Environ Microbiol 59(8):2465–2473.

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

Varma A, Palsson B. 1993b. Metabolic capacities of *Escherichia coli* II. Optimal growth patterns. J Theor Biol 165:503–522.

Varma A, Palsson B. 1995. Parametric sensitivity of stoichiometric flux balance models applied to wild-type *Escherichia coli* metabolism. Biotechnol Bioeng 45:69–79.

Yang YT, San KY, Bennett GN. 1999. Redistribution of metabolic fluxes in *Escherichia coli* with fermentative lactate dehydrogenase overexpression and deletion. Metab Eng 1(2):141–152.