Available online at www.sciencedirect.com
SCIENCE DIRECT
Metabolic Engineering 8 (2006) 338–352
METABOLIC ENGINEERING
www.elsevier.com/locate/ymben

The fractional contributions of elementary modes to the metabolism of *Escherichia coli* and their estimation from reaction entropies ☆

Aaron P. Wlaschin, Cong T. Trinh, Ross Carlson[1], Friedrich Srienc*

240 Gortner Laboratory, Department of Chemical Engineering and Materials Science, and BioTechnology Institute, University of Minnesota, 1479 Gortner Avenue, St. Paul, MN 55455/55108, USA

Received 18 November 2005; received in revised form 4 January 2006; accepted 31 January 2006
Available online 3 April 2006

Abstract

The metabolism of a cell can be viewed as a weighted sum of elementary modes. Due to the multiplicity of modes the identification of the individual weights represents a non-trivial problem. To enable the determination of weighting factors we have identified and implemented two gene deletions in combination with defined growth conditions that limit the metabolism from 4374 original elementary modes to 24 elementary modes for a non-PHB synthesizing control and 40 modes for a PHB synthesizing strain. These remaining modes can be further grouped into five families that have the same overall stoichiometry. Thus, the complexity of the problem is significantly reduced, and weighting factors for each family of modes could be determined from the measurement of accumulation rates of metabolites. Moreover, it is shown that individual weights are inversely correlated with the entropy generated by the operation of the used pathways defined in elementary modes. This suggests that evolution developed cellular regulatory patterns that permit diversity of pathways while favoring efficient pathways with low entropy generation. Furthermore, such correlation provides a rational way of estimating metabolic fluxes based on the thermodynamic properties of elementary modes. This is demonstrated with an example in which experimentally determined, intracellular fluxes are shown to be highly correlated with fluxes computed based on elementary modes and reaction entropies. The analysis suggests that the set of elementary modes can be interpreted analogous to a metabolic ensemble of quantum states of a macroscopic system.

Keywords: Elementary mode analysis; Poly (R)-3-hydroxybutyric acid; PHB; 3-hydroxybutryic acid; 3HB; Anaerobic metabolism; Metabolic modeling; *Escherichia coli* metabolism; Weighting factors; Statistical thermodynamics; Entropy

1. Introduction

Creating an in silico model of the metabolism of *Escherichia coli* has proved an attractive and elusive target. A large volume of work has been developed generating intricate metabolic networks linking enzymes, metabolites, and stoichiometry (Pramanik and Keasling, 1998; Aristidou et al., 1999; Lee and Papoutsakis, 1999; Edwards and Palsson, 2000; Burgard and Maranas, 2001; Carlson and Srienc, 2004a, b; Papin et al., 2004). Several groups have taken different approaches to processing this information. One of the most promising tools emerging for metabolic network investigation is elementary mode analysis (EMA). EMA reduces the complex metabolism of a cell into all possible, unique, non-divisible pathways for a network (Schuster et al., 1994, 2000; Heinrich and Schuster, 1996; Pfeiffer et al., 1999). An elementary mode represents a unique combination of stoichiometrically balanced enzyme catalyzed reactions within a cellular metabolism. Under steady-state conditions, the metabolic fluxes of an organism can be expressed as a non-negative, linear, weighted combination of elementary modes. Ultimately, one would like to identify the weighting factors to determine the fractional contribution of each elementary mode to the overall metabolism. The quantification of these weighting factors is difficult, if not impossible, since usually not

☆ Presented at the AIChE Annual Meeting, paper 153 d, 30 October–4 November 2005, Cincinnati, OH, USA.

* Corresponding author.
E-mail address: srienc@umn.edu (F. Srienc).
[1] Current address: Center for Biofilm Engineering, Montana State University, Bozeman, MT 59717, USA.

1096-7176/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2006.01.007

enough experimental information is available to explicitly solve for each weighting factor. Therefore, we have reduced the complexity of the metabolic network by reducing the number of possible elementary modes, which limits the quantity of unknown variables resulting in a solvable problem.

Early work in this field has focused on utilizing EMA to examine the central metabolic pathways (Dandekar et al., 1999). This soon expanded to models of the entire cell metabolism (Stelling et al., 2002). EMA can also be used as a guide to improve the production of desired metabolites. Some early success was established by improving the production of aromatic pathways in *E. coli* (Liao et al., 1996) and, more recently, by predicting the possibility of anaerobic PHB synthesis in *Saccharomyces cerevisiae* (Carlson et al., 2002), in suggesting improvements for recombinant protein production in bacteria (Vijayasankaran et al., 2005), and in the identification of the most efficient pathways for cell growth under different degrees of oxygen limitation (Carlson and Srienc, 2004a, b). Trinh verified the predictions made by Carlson with inverse engineering of the *E. coli* genome (Trinh et al., 2006). Recently an approach has been described to determine the contributions of individual elementary modes to the overall metabolism of *Lactobacillus rhamnosus* (Poolman et al., 2004). However, this work resulted in an underdefined system, with more unknowns than unique equations and, therefore, the individual contributions could not be precisely determined. Instead of trying to determine rates for an underdetermined system we used EMA as a guide to simplify the metabolism by reducing the number of possible elementary modes, making the set of equations solvable.

The system chosen for analysis is the transgenic, anaerobic production of poly (R)-3-hydroxybutyric acid (PHB) in *E. coli*. PHB belongs to a group of naturally occurring, biodegradable polyesters, known as polyhydroxyalkanoates (PHA). Their material properties make them a potential alternative to some petroleum-based thermoplastics (Sudesh et al., 2000). PHA is naturally synthesized in a wide variety of bacteria as a reserve compound for carbon and energy (Sipkema et al., 2000). Recombinant *E. coli* production of PHB has been extensively studied (Schubert et al., 1988; Kidwell et al., 1995; Lee et al., 1995; Choi et al., 1998; Wang and Lee, 1998; Lee and Choi, 2001; van Wegen et al., 2001). *E. coli* is a facultative anaerobe whose metabolism can also be adapted for anaerobic production of PHB (Carlson et al., 2005). Anaerobic production of PHB is attractive as an industrial process if the yields are not compromised (van Wegen et al., 1998).

EMA provides powerful insight about the intricacy of internal relationships within the cell metabolism. The significance of each elementary mode is enhanced when it can be coupled with its weighting factor. This combination reveals the fractional contribution of each elementary mode within the overall metabolism and permits computation of fluxes through each individual reaction. This allows identification of pathways that can be eliminated to improve production, and also permits prediction of the effects on the overall metabolism when these pathways are eliminated. This approach was used as a guide to improve the yields of PHB, and to provide more information about the effects of PHB production on the metabolism of a cell.

2. Theory

The metabolic network of a cell can be expressed in the classical, chemical reaction equation (Roels, 1983):

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

where $\mathbf{R}$ is a column vector comprised of net metabolite fluxes, $\boldsymbol{\alpha}$ is the stoichiometric matrix of all reactions and metabolites in the network, and $\mathbf{M}$ is a column velocity vector comprised of the fluxes through each reaction. This relationship is used in Metabolic Flux Analysis (MFA) to determine $\mathbf{M}$ from measured values of $\mathbf{R}$. Using the known stoichiometry of $\boldsymbol{\alpha}$, one can solve for $\mathbf{M}$, but this approach is dependent on having enough measurable metabolite fluxes to solve all the implied equations in (1). The approach of lumping reactions into groups shares a similar philosophy with extreme pathway analysis, however extreme pathways represent a subset within the set of identifiable elementary modes (Klamt and Stelling, 2003).

In contrast to MFA, in EMA the overall fluxes through the metabolic pathways described in the term $\mathbf{M}$, can be described as a weighted, linear combination of reactions specified in i elementary modes.

$$\mathbf{R} = \boldsymbol{\alpha} \sum_i \mathbf{E}_i c_i, \tag{2}$$

where c_i is a scalar representing the weighting factor that is the fractional contribution of each elementary mode $\mathbf{E}_i$ to the overall metabolic flux $\mathbf{M}$. Thus the sum of all the weighting factors is equal to one. This expression can be written in matrix notation as

$$\mathbf{R} = \boldsymbol{\alpha}\mathbf{E}\mathbf{c}, \tag{3}$$

where $\mathbf{E}$ represents a matrix with columns representing the individual elementary modes and $\mathbf{c}$ is a vector of the corresponding weights. The linear algebra dot product of $\boldsymbol{\alpha}$ with $\mathbf{E}$ results in

$$\mathbf{R} = \mathbf{N}\mathbf{c}. \tag{4}$$

Now $\mathbf{N}$ represents a matrix of net accumulation rates of metabolites with columns specifying net rates derived from each individual elementary mode. Values in $\mathbf{N}$ for internal metabolites are always zero. For external metabolites they usually are non-zero values but they can be also zero such as in the case of elementary modes that represent futile cycles. Identical columns in $\mathbf{N}$ indicate that the corresponding elementary modes have the same overall stoichiometry for the measurable, external metabolites. These modes can be grouped into a family of modes represented by a single column in a reduced net rate matrix $\mathbf{N}'$. Their

corresponding weighting factors can be added to result in a single weighting factor representing the entire family. This results in a reduced vector $\mathbf{c}'$ of weighting factors in the expression for the net accumulation rates of external metabolites:

$$\mathbf{R} = \mathbf{N}'\mathbf{c}'. \quad (5)$$

The advantage of Eq. (5) is that it consists only of relationships that are linearly independent. Therefore, the corresponding weighting factors can be determined from the measured accumulation rates of metabolites if the number of relationships equals or exceeds the number of unknown weighting factors. A further advantage is that the number of unknown weighting factors is reduced since a family mode represents several elementary modes. The family weighting factors represent an upper limit for possible fractional contributions of individual elementary modes in a family. This can be easily seen if one considers a case in which all elementary modes in a family, except one, do not contribute at all to the overall flux. In such case the entire family weight rests on the single, contributing elementary mode.

3. Materials and methods

3.1. Experimental strains

All experiments used a K12 *E. coli* strain, MG1655 (Blattner et al., 1997). Transforming *E. coli* with plasmid pPT500 generated a PHB synthesizing strain (Jackson and Srienc, 1994; Eschenlauer et al., 1996). Strains harboring this plasmid are referred to as PHB+. Plasmid pPT500 contains the native, three gene *Ralstonia eutropha* PHB operon from pAeT41 ligated into the pCR Blunt vector (Invitrogen, Carlsbad, CA) (Peoples and Sinskey, 1989). The operon is efficiently expressed in *E. coli* by the native *R. eutropha* operon promoter which resembles the *E. coli* σ^{70} system (Steinbuchel, 1991). A negative control was generated by creating a non-PHB producing plasmid. The control plasmid, pCR-KT, was constructed by ligating the *R. eutropha* beta-ketothiolase gene, without a promoter, into the pCR Blunt vector. Included on each plasmid is the kanamycin resistance gene.

3.2. Gene knockout strain construction

Strains containing the genetic knockouts *ldh* and *frdA* were obtained from the library of *E. coli* knockouts "KO-V1-sg" deposited in the AGAC facility at the University of Minnesota. This library was generated from clones of an original library created by Mori (Mori et al., 2000). The knockouts were transferred into the MG1655 strain. The transfer and accumulation of the knockouts in a single strain was accomplished using kanamycin resistance and the λ Red phage system (Datsenko and Wanner, 2000). Pairs of oligonucleotides, from Integrated DNA Technologies, Inc. (Coralville, IA), designed inside the targeted knockout gene and to regions outside the targeted knockout gene, (Table 1) were used in conjunction with PCR and gel electrophoresis to verify the elimination of the target gene and investigate the possibility of the appearance of the target gene somewhere else within the genome (results not shown). PCR was done with kits purchased from GE Healthcare (Piscataway, NJ). Plasmids were isolated with commercial preparation kits (Qiagen, Hilden, Germany). The three strains generated were: MG1655 Δ*frdA,* MG1655 Δ*ldh*, and MG1655 Δ*frdA* Δ*ldh*. These strains were then transformed with either the PHB+ or the PHB− plasmid.

3.3. Inoculum and storage conditions

The masterstocks of each strain used in this study were stored at −80 °C in a 20% glycerol solution. Inoculum for each reactor was generated by thawing a 1.5 ml sample from the masterstock and incubating it for 24 h at 30 °C and 150 rpm in M9 medium (1.0 g/L NH_4Cl, 0.5 g/L NaCl, 6.0 g/L Na_2HPO_4 3.0 g/L KH_2PO_4, 3 mg/L $CaCl_2$, 1 ml/L 1 M $MgSO_4$ (Sigma, St. Louis, MO)). One ml of the concentrated trace minerals and metals solution (0.15 g/L H_3BO_4, 0.065 g/L $CoSO_4$, 0.05 g/L $ZnSO_4$–$7H_2O$, 0.015 g/L $MnCl_2$–$4H_2O$, 0.015 g/L $NaMo_4$–$2H_2O$, 0.01 g/L $NiCl_2$–6-H_2O, 0.005 g/L $CuSO_4$–$5H_2O$, 3.0 g/L Fe(NH_4)-citrate (Sigma, St. Louis, MO). was added to every liter of M9 prepared. Glucose was supplied at a concentration of 40 g/L and thiamine was added 5 mg/L (Sigma, St. Louis, MO) (Kawasaki et al., 1968). Selection pressure was applied by the addition of 100 μg/L kanamycin. To create microaerobic conditions for the inoculum and reduce the lag phase

Table 1
Primers used to verify the knockout mutations on the genome

Gene	Upstream	Downstream
Outside primers		
Ldh	5′-CGC AAC AAA CGC GGC TAC-3′	5′-CGG CTT TAT ATT TAC CCA-3′
FrdA	5′-CGG TAA TTA ATA AGG GCG AGA GCG-3′	5′-CTC CAG TTT TTG ACA AGG GC-3′
Inside primers		
Ldh	5′-TAC CCA ACG AAC CAA TTT TC-3′	5′-GCT GGA AGA GCT GAA AAA GC-3′
FrdA	5′-TTA CGT GCC ATT GCG GAG TG-3′	5′-TCA CGA TAC AGT AGC GGG TG-3′

after inoculation in the bioreactor, the inoculum was grown in a sterile polypropylene tube filled completely with medium to minimize headspace. A pinhole was punched in the cap and covered with tape. This was done to prevent a buildup of pressure due to carbon dioxide gas.

3.4. Bioreactor conditions

The fermentations used the same medium and conditions as the inoculum. Anaerobic conditions were established by sparging the bioreactors overnight with nitrogen gas to purge dissolved oxygen from the medium. The absence of oxygen was confirmed by DO probe (Mettler-Toledo, Columbus, OH) and by off-gas mass spectrometry, ThermoOnix Prima δB mass spectrometer (Houston, TX). After inoculation, a constant flow rate of 0.5 L/min of nitrogen gas was sparged to provide positive pressure and maintain anaerobic conditions. All cultures were maintained at 30 °C, with the pH controlled at 7.0 using 4 M NaOH and with an agitation rate of 300 RPM.

3.5. Analytical techniques

Samples were regularly taken from the bioreactors. They were quickly separated by centrifugation to isolate the biomass from the supernatant. The biomass was dried overnight at 110 °C and weighed. The supernatant was frozen at −20 °C until the experiment was finished. PHB content of the samples was prepared by propanolysis and detected using the gas chromatography method (Riis and Mai, 1988). Samples were run on a Hewlett-Packard 5890A gas chromatograph (Palo Alto, CA) with a DB-WAX 30W column and were analyzed with a flame ionization detector (FID). Glucose concentrations were determined by absorption at 340 nm using hexokinase enzyme kits (Sigma, St. Louis, MO). Ethanol and acetate concentrations were determined with GC on the same gas chromatograph and detector, but using a DB-FFAP capillary column from the manufacturer. One ml of the supernatant was syringe filtered and acidified. An internal standard was simultaneously added by including 0.1 ml of 1 N HCl, in 10 g/L 1-propanol (Sigma, St. Louis, MO)). The by-products lactate, succinate, and 3-hydroxybutyrate were analyzed as described (Vollbrecht et al., 1978) using the same gas chromatograph and detector, but with a DB-WAX 30W column. CO_2 measurements were obtained from the fermentor off-gas mass spectrometer described earlier.

3.6. Elementary mode analysis

A metabolic map for *E. coli* was generated using biochemical pathways reported in literature. A detailed description of the model has previously been published (Carlson and Srienc, 2004a). Several modifications were made to adapt the model to PHB production. These additional adjustments have also been previously described (Carlson et al., 2005). The following two additional reactions were included in the metabolic network to account for the free monomer, 3-hydroxybutyric acid (3HB). The first reaction, R99, accounts for the synthesis of 3HB from acetyl-CoA. 3HB was found in the supernatant so an additional reaction was required to represent the secretion of this metabolite to an external pool.

$$\mathrm{R99}: 2\,\mathrm{ACETYL\_CoA} + \mathrm{NADH} = 3\mathrm{HB} + 2\mathrm{CoASH} + \mathrm{NAD},$$

$$\mathrm{R101}: 3\mathrm{HB} = 3\mathrm{HB}\ (\mathrm{external}).$$

These reactions are defined according to the format described in the original metabolism model. The single reaction, R99, is based on the two of the three *R. eutropha* PHB pathway genes, the enzymes beta-thiolase, and reductase. The *R. eutropha* reductase is NADPH dependent. The cleavage of the 3-hydroxybutyric–CoA complex is likely accomplished by a native thioesterase, or a combination of steps involving the native phosphotransacetylase and acetate kinase. The final model is comprised of 51 total reactions (22 reversible and 29 irreversible) reactions involving 50 metabolites (13 external and 37 internal). The model was compiled using the publicly available METATOOL program, version meta4.3_double (Schuster et al., 1994; Pfeiffer et al., 1999).

4. Results

The combined effect of two gene knockouts, anaerobic growth conditions, and a defined growth medium significantly reduced the number of elementary modes available in the PHB production phase from 4374 to 40. A summary of the number of elementary modes remaining after implementing each of these conditions is given in Table 2. In a non PHB-producing metabolic network, the number of elementary modes was reduced to 24. The EMA guided strategy was directed towards keeping modes with high PHB yield and reducing the number of families, while still

Table 2
Number of elementary modes available for PHB producing *E. coli* metabolism after each genetic knockout or environmental restriction

Genetic or environmental conditions	Number of available modes
Total modes	4374
Total anaerobic modes	614
Anaerobic, production phase	243
Anaerobic, growth phase	371
Anaerobic, no lactate, production phase	191
Anaerobic, no succinate, production phase	48
Anaerobic, no lactate, no succinate, production phase	40

The growth phase refers to modes that are linked to the formation of biomass. The production phase refers to modes that do not produce biomass and synthesize only PHB and fermentation products. They are independently available during biomass formation, but are also available during periods of nutrient limitation.

maintaining enough measurable metabolite fluxes to uniquely solve for family weighting factors. The first target was the removal of lactate from the metabolite pool. This was experimentally accomplished by knocking out the native lactate dehydrogenase, *ldh*. The second knockout mutation was the elimination of the ability to produce succinate from fumarate. Under anaerobic conditions, the TCA cycle is not complete, and only operates as a branched system (Clark, 1989). Thus the elimination of fumarate reductase, *frdA*, removes the capability to produce succinate from fumarate from the TCA cycle (Maklashina et al., 1998). The remaining modes were grouped into five families based on their identical overall stoichiometry of remaining secreted metabolites.

4.1. Growth characteristics

All knockout strains generated were able to grow under anaerobic conditions. The medium was designed to limit biomass on nitrogen. Residual cell dry weight (RCDW) time profiles are given in Fig. 1. RCDW was calculated by subtracting the mass of PHB from the total cell dry weight. In addition to measuring intracellular PHB concentrations, data was collected on secreted ethanol, acetate and secreted 3HB. Lactate and succinate concentrations were also determined however, succinate and lactate were not detected when either *frdA* or *ldh* were deleted. This is a secondary verification for the removal of these gene functions from the cellular metabolism. CO_2 and H_2 were recorded in the fermentor exhaust gas for all the experiments.

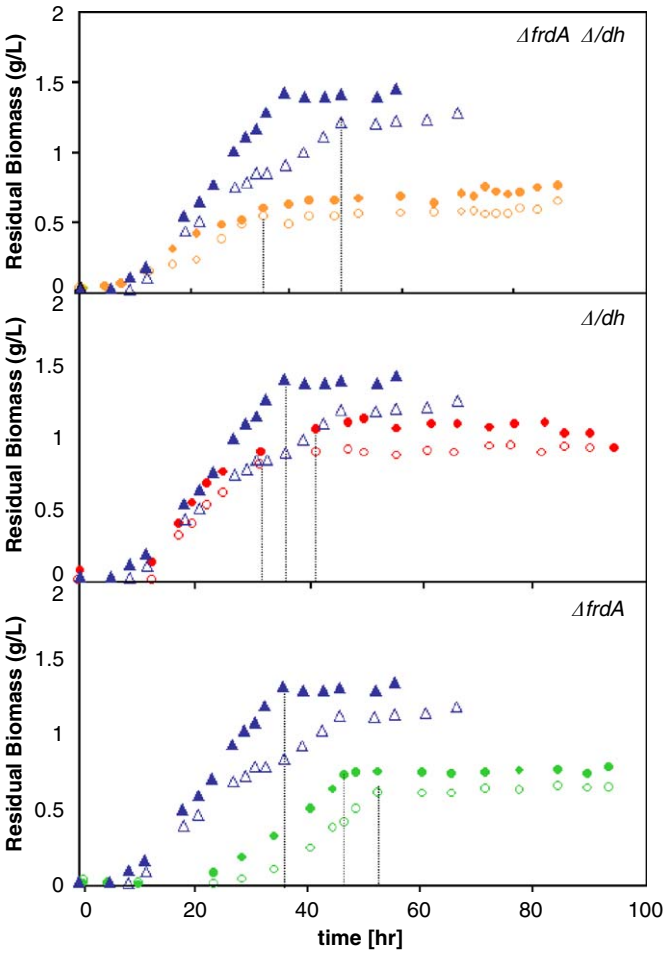

Fig. 1. Measured residual biomass for each strain. Open circles refer to the PHB+ strain, closed circles are the PHB− control. The time profiles for the wild type strain (PHB+ open triangles, PHB− closed triangles) are shown in all graphs for comparison. The dashed lines refer to the start of production phase for each culture.

Based on the measurements of all the metabolites and the composition of the exhaust gas, a complete carbon balance was calculated for each sample from the production phase. The maximum error calculated was $\pm 7\%$ compared to the initial amount of carbon. Most of the samples had an error of $\pm 3\%$. RCDW was estimated to be 50% carbon for anaerobically grown *E. coli* in stationary phase (Heldal et al., 1985; Neidhardt, 1996). The accountability of formate provided a challenge. It was necessary to separate formate from carbon dioxide to apply the EMA and to solve for weighting factors. Formate can quickly react with ions in the water to decompose into carbon dioxide and water, thus obscuring the separate amount of each chemical (Talu and Diyamandoglu, 2004). This estimation was also needed for the thermodynamic analysis to account of the difference in the amounts of heat and free energy required to synthesize the different compounds. To separate the free carbon dioxide generated by the cellular metabolism from the carbon dioxide generated from decomposed formate, an estimate was developed according to the stoichiometry of each measured metabolite.

$$mole_{\text{acetate}} + mole_{\text{ethanol}} + 2\,mole_{\text{PHB}} + 2\,mole_{\text{3HB}} = mole_{\text{formate}}. \quad (6)$$

This equation is not a chemical reaction; rather it is assumed in (6) that every mole of formate is formed via pyruvate formate-lyase, *pfl*, during the synthesis of one mole of acetyl-CoA. The sum of moles of acetate and moles of ethanol approximates the total amount of acetyl-CoA. In addition, each mole of 3HB or mole of PHB is synthesized from two moles of acetyl-CoA. The amount of carbon dioxide that could be formed from the concentration of formate calculated from (6) was below the concentration detected by the exhaust gas mass spectrometer. Reevaluation based on weighting factors and elementary modes reveals approximately all of the flow of carbon occurs through *pfl*. Further refinements with this information yielded no significant change in the final results.

4.2. Effect of mutations on metabolite secretion and PHB accumulation

The gene knockouts targeting the secretion of lactate and succinate, also affected the production of other fermentation metabolites. The specific amount of acetate was decreased in the Δ*frdA* Δ*ldh* strain when compared to the

wild type for PHB+ experiments (Fig. 2). The reverse was true for the PHB− experiments (Fig. 3). The difference in the PHB+ results may be due to the competition for acetyl-CoA, a common precursor to both PHB and acetate. The synthesis rate of ethanol was not changed significantly by the mutations or the production of PHB. The PHB+ and PHB− versions of the Δ*frdA* Δ*ldh* strains utilized different modes to produce ethanol, yet the observed concentrations were similar.

It was found that the single *frdA* mutation removed over 80% of the modes when compared to a PHB synthesizing metabolic network, yet the elimination of a majority of modes lowered the yield of PHB on glucose when compared to a wildtype PHB synthesizing strain. Although the *frdA* mutation removed some of the available PHB producing modes, the highest PHB yielding modes remained. The single *ldh* mutation led to similar results, although fewer modes were removed. The moles of HB (moles of 3HB plus moles of PHB) yield measured from the wild type HB synthesizing strain was 0.130 mole HB/mole Glucose. As seen in Fig. 4, the measured yield of the Δ*ldh* strain was 0.106 mole HB/mole Glucose. The Δ*frdA* strain had a similar yield of 0.103 mole HB/mole glucose. The *ldh* or the *frdA* knockouts affect the ability to secrete reduced metabolites and to regenerate the NAD cofactors. The data shows that in the *ldh* knockout lactate secretion is eliminated. But the cells compensate for this by increasing succinate secretion. The opposite is true for the mutant that contains only the *frdA* deletion. In both of these mutants the PHB synthesis apparently is not used as an additional electron sink to compensate for the mutations since the yield of PHB is not significantly affected. The elimination of both *frdA* and *ldh* is apparently needed to force cells to use PHB as the electron sink for recycling reducing equivalents.

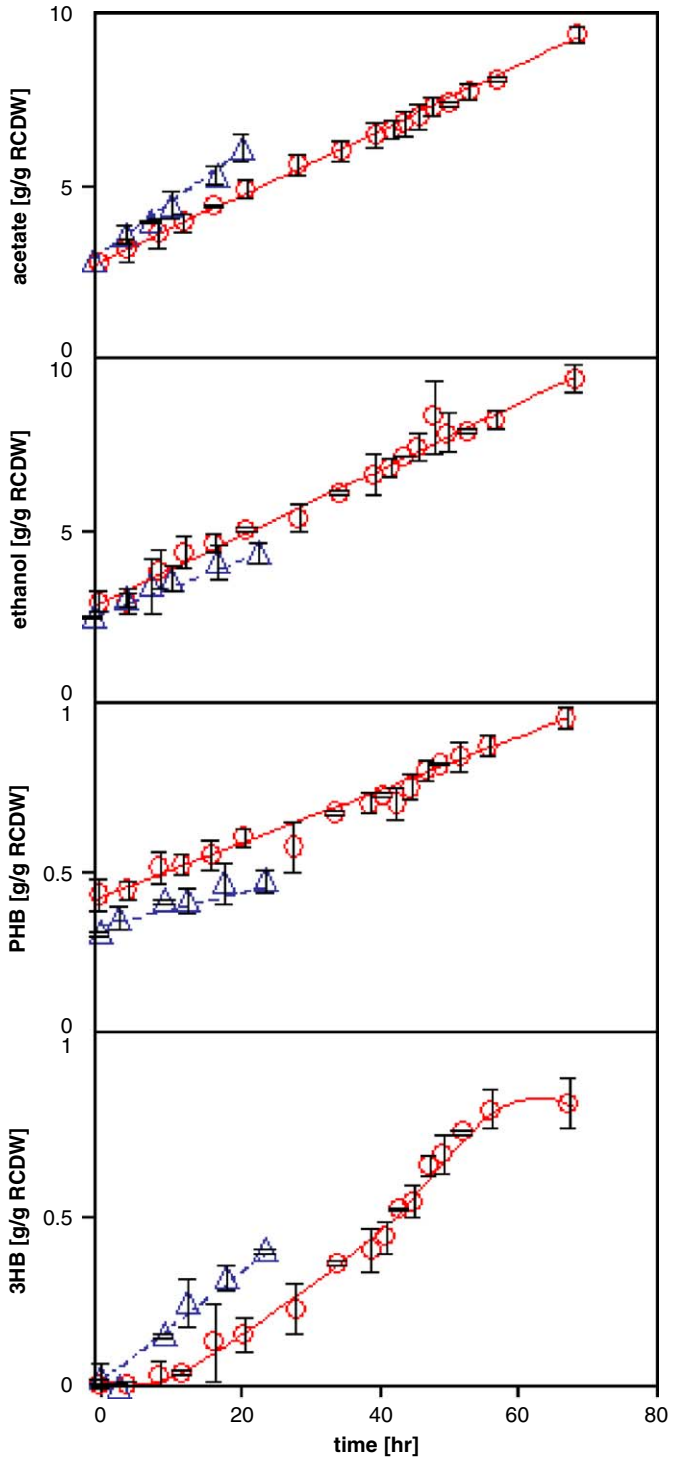

Fig. 2. Specific concentrations of secreted fermentation products during production phase for PHB+ strains of Δ*frdA* Δ*ldh* (red open circles), and PHB+ wild type (blue open triangles).

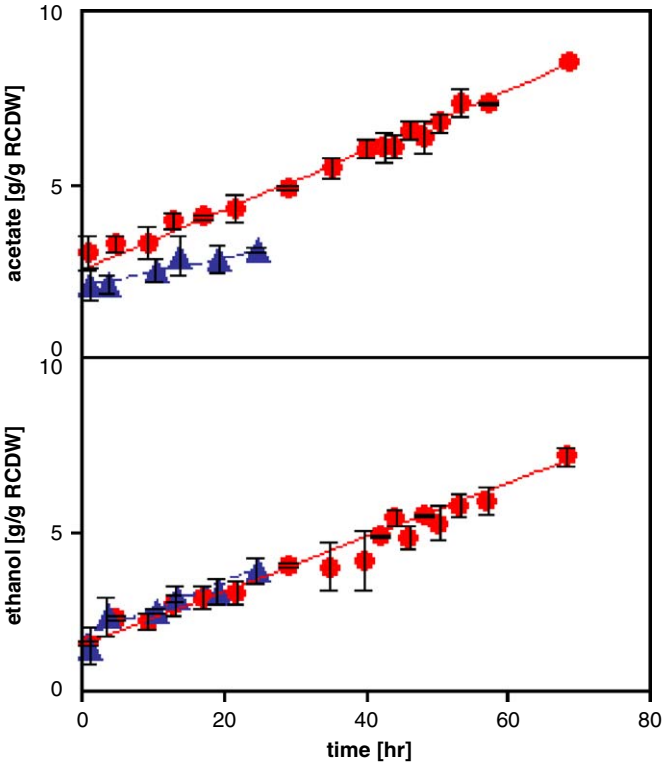

Fig. 3. Specific concentrations of secreted fermentation products during production phase for PHB− Δ*frdA* Δ*ldh* (red closed circles), and PHB− wild type (blue closed triangles).

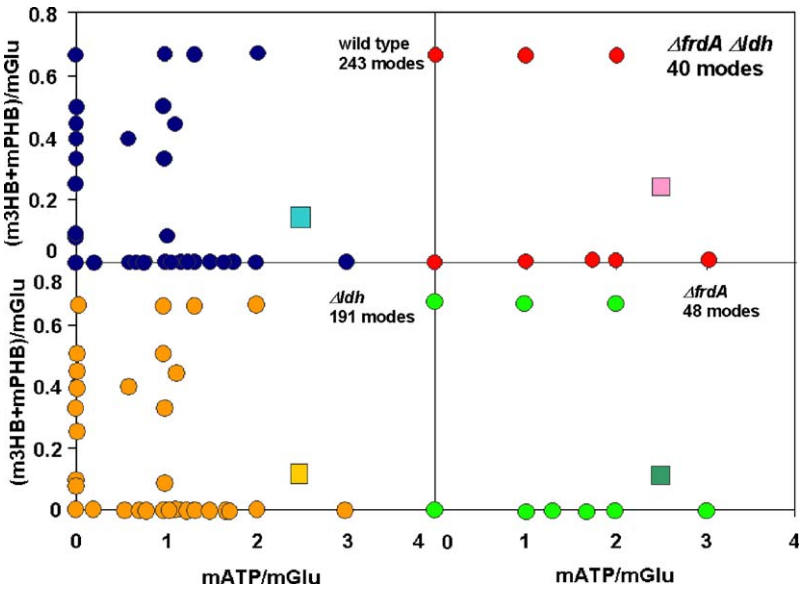

Fig. 4. Comparison between available elementary modes and measured yields. The circles represent the moles HB (moles 3HB plus moles PHB) consumed compared to the moles of unused ATP generated, normalized by the moles of glucose consumed. One point can represent several modes that have the same yields. The squares represent the yields from each respective experiment. A scaled, linear combination of the values from the elementary mode positions will describe the position of the square. The values for the maintenance energy requirements were developed from data developed in chemostats under similar conditions published by Mainzer according to the method originally described by Pirt (Pirt, 1975; Mainzer and Hempfling, 1976).

The double knockout strain also secreted the PHB monomer, 3HB. The yield in moles of 3HB and PHB was pooled in this analysis to determine the flux directed toward the PHB pathway. The appearance of 3HB is likely due to the slower rate of the final synthase step in the PHB synthesis sequence. The polymerization reaction may not have been able to utilize all of the free monomer produced, leading to the overproduction of 3-hydroxybutyric-CoA. This compound can be likely broken down by thioesterase enzymes in the cell, leading to the secretion of the 3HB. Normally, no step in the PHB synthesis pathway is considered limiting, however under the anaerobic conditions designed for the Δ*frdA* Δ*ldh* study, the cofactors acetyl-CoA and NADPH are likely present in abundance. This provides ideal conditions that increase the rate of the ketothiolase and reductase in the *R. eutropha* PHB synthesis pathway (Leaf and Srienc, 1998; van Wegen et al., 2001). The high PHB yielding modes are always an option for the cell to use, but the combination of *frdA* and *ldh* mutations eliminates all of the alternative lower PHB-yielding modes. The cellular metabolism is forced to direct more carbon into PHB production, and the only remaining PHB modes are the modes with the highest PHB yields. The wild type strain had an average yield of 0.130 moles of HB per mole of glucose consumed during production phase. The combined effect in the Δ*frdA* Δ*ldh* strain raised the yield by 75% to 0.230 mole HB/mole glucose. Fig. 4 describes these results. This is still well below the maximum theoretical yield of 0.66 mole HB/mole glucose, but represents a significant improvement over the yields in the wild type strain.

The stoichiometry of modes determines the coordinates of the data points in Fig. 4. The number of points expresses the number of available modes left in the metabolic network. The HB yields (squares in Fig. 4) were experimentally determined. The values for the maintenance energy requirements were taken from data developed in chemostats under similar conditions published by Mainzer (Mainzer and Hempfling, 1976) according to the method originally described by Pirt (Pirt, 1975).

4.3. Families of modes

The elementary mode model developed for this study identified 4374 overall modes. Unfortunately, solving for this many unknown weighting factors is essentially impossible. Metabolite balances, boundary conditions, and measurable fluxes do not provide enough unique conditions to make the system solvable. To reduce the number of unknown variables, the elementary modes were grouped into families based on their overall stoichiometry. This simplified the number of unknown weighting factors to 3547. A further reduction was obtained through a combination of gene knockouts, medium design and anaerobic growth conditions that limited this strain to 24 modes. The description of the available modes for the Δ*frdA* Δ*ldh* PHB− strain is given in Table 3. These resulting 24 modes were grouped into three families based on their identical, overall stoichiometry. Each family is comprised of eight unique elementary modes as indicated by the shaded sets of columns. For the Δ*frdA* Δ*ldh* PHB+ strain under the same conditions, 40 modes are available. The additional 16 modes are grouped into two additional families as shown in Table 4. Each member of a family utilizes a different combination of internal reactions, but the net accumulation rates of external metabolites and

Table 3
Stoichiometry of available modes for the Δ*frdA* Δ*ldh* strain

		Family 1								Family 2								Family 3							
Glucose uptake	R1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ethanol synthesis	R54, R90	1	1	1	1	1	1	1	1	1.71	1.71	1.71	1.71	1.71	1.71	1.71	1.71	2	2	2	2	2	2	2	2
Acetate synthesis	R55, R91	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Formate secretion	R96	2	2	2	2	2	2	2	2	1.71	1.71	1.71	1.71	1.71	1.71	1.71	1.71	0	0	0	0	0	0	0	0
PHB synthesis	R98, R100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3HB synthesis	R99, R101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Glycolysis	R2	1	1	1	1	1	1	1	1	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	1	1	1	1	1	1	1	1
	R3	4	3	1	1	1	1	1	1	0.71	0.71	2.43	1.71	0.71	0.71	0.71	0.71	1	1	1	1	3	1	1	2
	R4	3	2	0	0	0	0	0	0	0	0	1.71	1	0	0	0	0	0	0	0	0	2	0	0	1
	R5, R6*	1	1	1	1	1	1	1	1	0.71	0.71	0.71	0.71	0.71	0.71	0.71	0.71	1	1	1	1	1	1	1	1
	R7, R8	2	2	2	2	2	2	2	2	1.71	1.71	1.71	1.71	1.71	1.71	1.71	1.71	2	2	2	2	2	2	2	2
	R9	1	0	4	1	0	0	0	1	0	0.71	0.71	0	2.43	0.71	0	0	0	0	1	3	1	1	0	0
	RR9	0	0	3	0	0	1	0	0	0	0	0	0	1.71	0	0	0.50	0.5	0	0	2	0	0	0	0
Pentose phosphate pathway	R10	0	0	0	0	0	0	0	0	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0	0	0	0	0	0	0	0
	R11	0	0	0	0	0	0	0	0	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0	0	0	0	0	0	0	0
	R12, R13, R14, R15	0	0	0	0	0	0	0	0	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0	0	0	0	0	0	0	0
Pyruvate utilization	R20	2	2	2	2	2	2	2	2	1.71	1.71	1.71	1.71	1.71	1.71	1.71	1.71	0	0	0	0	0	0	0	0
	R21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	2	2	2	2	2	2
Anaplerotic reactions	R29	0	−1	0	0	−1	−2	−1	0	−0.71	0	0	−0.71	0	0	−0.71	−1.21	−1.5	−1	0	0	0	0	−1	−1
	R40	0	1	0	3	3	2	1	0	0.71	0	0	0.71	0	1.71	1.71	1.21	1.5	1	0	0	0	2	2	1
	R41	0	1	0	0	1	2	1	0	0.71	0	0	0.71	0	0	0.71	1.21	1.5	1	0	0	0	0	1	1
	R42	0	0	0	3	2	0	0	0	0	0	0	0	0	1.71	1	0	0	0	0	0	0	2	1	0
CO_2 release	R97	0	0	0	0	0	0	0	0	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86	2	2	2	2	2	2	2	2
ATP production	R82	0	0	0	0	0	0	2	3	1	1.71	0	0	0	0	0	0	0	1	2	0	0	0	0	0

Families 1, 2 and 3 are present in both the PHB− and PHB+ synthesizing strains. The flux of moles in R6 is opposite in the direction specified. R99 and R101 are specified in this text, R98 and R100 are defined in Carlson et al. (2005). The rest of the reactions are described in Carlson and Srienc (2004a). The dashed line separates the fluxes from external and internal metabolites. See the text for more details.

Table 4
Stoichiometry of available modes for the Δ*frdA* Δ*ldh* strain

		Family 4						Family 5								
Glucose uptake	R1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ethanol synthesis	R54, R90	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67
Acetate synthesis	R55, R91	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Formate secretion	R96	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
PHB synthesis	R98, R100	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0	0	0	0	0	0	0
3HB synthesis	R99, R101	0	0	0	0	0	0	0	0	0.67	0.67	0.67	0.67	0.67	0.67	0.67
Glycolysis	R2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	R3	1	3	1	1	2	1	1	1	1	1	3	1	1	2	1	1
	R4	0	2	0	0	1	0	0	0	0	0	2	0	0	1	0	0
	R5, R6*	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	R7, R8	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	R9	3	1	1	0	0	0	1	0	0	3	1	1	0	0	0	1
	RR9	2	0	0	0	0	0	0	0.5	0.5	2	0	0	0	0	0	0
Pentose phosphate pathway	R10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	R11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	R12, R13, R14, R15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pyruvate utilization	R20	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
	R21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Anaplerotic reactions	R29	0	0	0	−1	−1	−1	0	−1.5	−1.5	0	0	0	−1	−1	−1	0
	R40	0	0	2	2	1	1	0	1.5	1.5	0	0	2	2	1	1	0
	R41	0	0	0	1	1	1	0	1.5	1.5	0	0	0	1	1	1	0
	R42	0	0	2	1	0	0	0	0	0	0	0	2	1	0	0	0
CO_2 release	R97	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ATP production	R82	0	0	0	0	0	1	2	0	0	0	0	0	0	0	1	2

Families 4 and 5 are only present in the PHB+ synthesizing strains. The flux of moles in R6 is opposite in the direction specified. R99 and R101 are specified in this text, R98 and R100 are defined in Carlson et al. (2005). The rest of the reactions are described in Carlson and Srienc (2004a). The dashed line separates the fluxes from external and internal metabolites. See the text for more details.

Table 5
These are the overall stoichiometries for each family. The characteristics refer to an identifiable attribute of each family

Family	Characteristics	Stoichiometry
1	Acetate synthesis	Glucose = Ethanol + acetate + 2 formate
2	Pentose phosphate utilizing	Glucose = 12/7 Ethanol + 6/7 CO_2 + 12/7 formate
3	Pyruvate dehydrogenase complex utilizing	Glucose = 2 Ethanol + 2 CO_2
4	PHB synthesis	Glucose = 2/3 Ethanol + 2 formate + 2/3 PHB
5	3HB synthesis	Glucose = 2/3 Ethanol + 2 formate + 2/3 3HB

Table 6
Measured and calculated metabolite flux vectors, **R** [mmole/g RCDW h], for PHB+ and PHB− systems

	PHB+ metabolite fluxes (mmole/g RCDW h)		PHB− metabolite fluxes (mmole/g RCDW h)	
	$\mathbf{R}_p$	$\mathbf{R}_{cp}$	$\mathbf{R}_c$	$\mathbf{R}_{cc}$
Glucose	−1.000	−0.958	−1.000	−0.946
Ethanol	0.933	0.955	0.999	1.027
Acetate	0.534	0.547	0.812	0.839
Formate	1.801	1.803	1.820	1.825
PHB	0.066	0.098	—	—
3HB	0.066	0.098	—	—

The measured flux vectors, $\mathbf{R}_p$ and $\mathbf{R}_c$, correspond to the PHB+ and PHB− strains of Δ*frdA* Δ*ldh* genotype. These values were determined from supernatants of collected samples. They were scaled to the specific consumption rate of glucose. The calculated metabolite flux vectors, $\mathbf{R}_{cp}$ and $\mathbf{R}_{cc}$ are the product of the weighting factor and stoichiometry of the respective family for the PHB+ and PHB− models. The order of metabolites in the column corresponds to the order they appear in the elementary mode matrix described in the text.

Table 7
The weighting factors vector, $\mathbf{c}'$, for each family needed to define the metabolism of the PHB+, $\mathbf{c}'_p$, and PHB−, $\mathbf{c}'_c$, strains during production phase

Family	Weighting factors			
	PHB+, $\mathbf{c}'_p$	PHB−, $\mathbf{c}'_c$	PHB+, $\mathbf{c}'_{tp}$	PHB−, $\mathbf{c}'_{tc}$
1	0.547	0.839	0.281	0.915
2	0.069	0.086	0.075	0.069
3	0.046	0.02	0.036	0.016
4	0.148	—	0.330	—
5	0.148	—	0.277	—

The values for $\mathbf{c}'_{tp}$, and $\mathbf{c}'_{tc}$ are the weighting factors calculated from entropies of the families. See the text for details on the calculations.

consequently the overall metabolic stoichiometry is the same. The metabolic stoichiometry of each family is given in Table 5.

The PHB+ model has five unknowns (weighting factors representing each of the five families), and a total of seven equations (for each of the metabolites measured, 3HB, PHB, acetate, ethanol, formate, carbon dioxide and the substrate glucose). For further analysis, the carbon dioxide consideration was removed from the calculation. This was done due to concerns about the ability to completely flush the line between the reactor headspace and off gas mass spectrometer with the low flow rate of sparging nitrogen gas resulting in inaccurate data.

An overdetermined system does not guarantee a logical solution set of weighting factors (or even a solution set at all), instead we rely on a least-square method to solve the system of equations. An error value was calculated as the difference between the measured metabolic fluxes and calculated metabolic fluxes Table 6. This error was minimized. The error of the final solution set is several orders of magnitude below that of measured values, thus ensuring this is at least a very feasible, if not the best solution. The obtained weights are listed in Table 7.

With the weighting factors defined one can compute the intracellular metabolic fluxes as weighted averages of elementary modes (Fig. 5). This is straightforward for elementary mode reactions that are identical in each family since in this case the same family weighting factor applies to all modes. In cases where fluxes vary between elementary modes of a family a defined value cannot be assigned to the overall reaction fluxes on the basis of only family weighting factors. But in such cases a flux range can be calculated from the elementary mode fluxes and the family weighting factors. These are reactions enclosed in the shaded boxes of Fig. 5. The flux ranges for these reactions are specified in Table 8. One should note that experimental determination of these fluxes, for instance with labeling experiments, would permit further discrimination of weighting factors associated with elementary modes within a family of modes. But such experiments have not been done in this work.

4.4. Thermodynamics of elementary modes

The individual weighting factors can be formally viewed as probabilities that a given elementary mode is used within a functioning metabolism. We explored a possible relationship between the use of specific pathways as expressed by the weighting factors, and the thermodynamic properties of the reaction sequence specified by these pathways. Under non-growth conditions, the metabolic network of a cell collects the chemical bond energy from high-enthalpy, low-entropy ordered substrates, by converting them into low-enthalpy, high-entropy products. This constant flux of products and substrates can be modeled as an open system

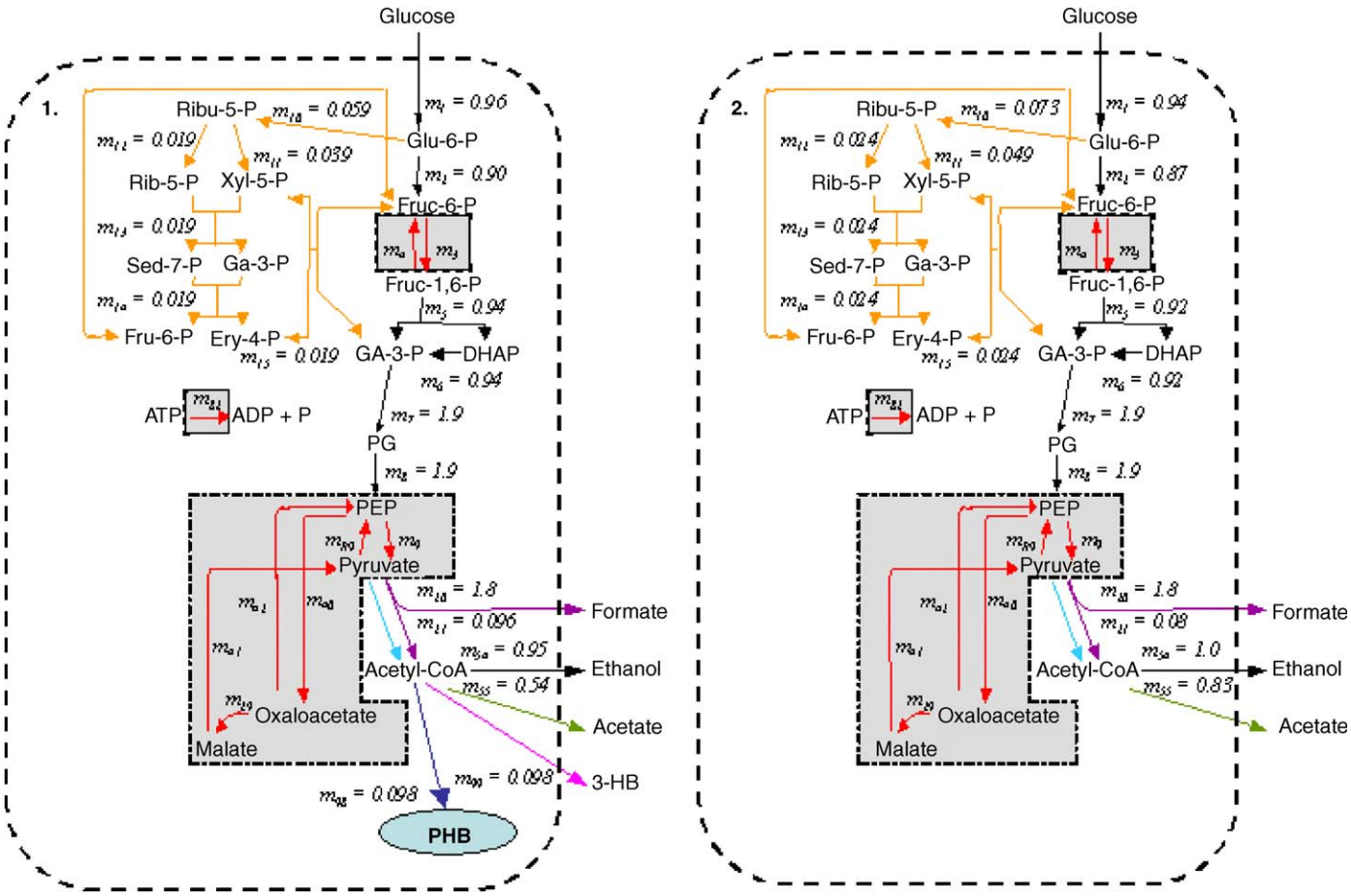

Fig. 5. Metabolic flux maps for Δ*frdA* Δ*ldh* strains for the PHB+ (1.) and PHB− control (2.) The fluxes, m [mmole/g RCDW h], were determined from the combination of the family weighting factors and the stoichiometry of elementary modes. The known fluxes are given; the unknown fluxes are discussed in Table 8. The label of each flux corresponds to the reaction number given in Tables 3 and 4. The color of the arrows corresponds to their use by different families. Fluxes utilized by all modes are black, fluxes unique to family #1 are green, family #2 are light orange, family #3 are light blue, family #4 are dark blue, and family #5 are pink. Families #1, #2, #4 and #5 utilize some pathways, these fluxes are purple. For the PHB− metabolic map, families 4 and 5 are not present. The shaded boxed areas represent fluxes that would not be able to be calculated with MFA. Combinations of individual family members utilize the red fluxes within these boxes.

Table 8
Range of fluxes, m [mmole/g RCDW h] that cannot be explicitly defined for Δ*frdA* Δ*ldh*

	m3	m4	m9	mr9	m40	m41/m29	m42	m82
PHB+								
Maximum	3.39	2.45	3.39	2.45	2.45	1.69	2.45	2.45
Minimum	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PHB−								
Maximum	3.62	2.70	3.62	2.70	2.70	1.81	2.70	2.70
Minimum	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The fluxes were determined by combining the maximum and minimum stoichiometry of each family member with the appropriate weighting factor. The label of each flux corresponds to the reaction number given in Tables 2 and 3.

in classical thermodynamics. To measure the energy of this process, we can use the degree of reduction method described by Roels, and updated by Sandler (Roels, 1983; Sandler and Orbey, 1991). They provide a method for estimating the heat and the free energy of combustion for metabolites based on the structure of the molecules involved. This method was adapted to find the heat and the free energy values for 3HB and PHB by estimating needed correction factors (Sandler and Orbey, 1991) from a correlation between degrees of reduction and correction factors given for other, similar metabolites. This information was combined with the family reaction stoichiometry to estimate the enthalpy, Gibbs free energy and entropy for each family based reaction (Table 9).

Several relationships between the thermodynamic quantities and the weights were tested but the best fit was obtained in a plot of the reaction entropies of each family versus the natural log of weighting factors (Fig. 6). For the PHB− system the entropies were perfectly correlated with the weights of corresponding families. A similar correlation can be seen also for the PHB+ system but the fit is not as good.

The correlation suggests another method to estimate the weighting factors. If the correlation is valid the entropies

Table 9
Calculated Gibbs free energies, enthalpies, and entropies from reaction stoichiometries for the families in Tables 2 and 3 per mole of glucose consumed

Family	$\Delta G'^{\circ}$ (kJ/mole)	$\Delta H'^{\circ}$ (kJ/mole)	$\Delta S'$ (kJ/K mole)
1	−377.98	−307.01	0.2341
2	−376.25	−248.01	0.4230
3	−234.00	−69.02	0.5442
4	−377.67	−313.98	0.2101
5	−314.66	−242.98	0.2364

Degrees of reduction, heat and free energies of combustion and were calculated according to the method described in Sandler and Orbey (1991).

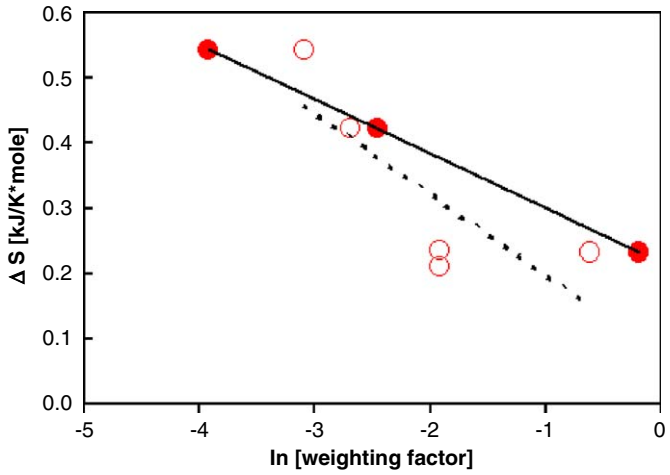

Fig. 6. Comparison of entropy for the PHB− (closed circles) and PHB+ (open circles). The weighting factors are for the Δ*frdA* Δ*ldh* metabolic network. The axes incorporate a logarithmic scale for the weighting factor. A linear relation is present for entropy for the PHB− system. This is indicated by the solid line with a slope of −0.083 [kJ/K mole ln(weighting factor)] and intercept of 0.22 [kJ/K mole]. The pattern for the PHB+ values is less defined as indicated by the poorly fitting dashed trend line. The line with the best fit for the PHB+ system has a slope of −0.12 [kJ/K mole ln(weighting factor)] and an intercept of 0.075 [kJ/K mole].

defined by the overall stoichiometry of n families or elementary modes can be related to the n unknown weighting factors in n relationships

$$\Delta S_j = a + b\,\ln(c_j) \quad j = 1,\ldots,n, \tag{7}$$

where a and b are constants defining the correlation, c_j are the weights and ΔS_j are the entropies generated per mole of glucose consumed for the jth elementary mode. Thus, the system consists of n equations and $n+2$ unknowns, and two more equations are needed to solve for a and b. One of them is the requirement that the sum of the weighting factors equals to one:

$$1 = \sum_j c_j. \tag{8}$$

If the overall metabolite fluxes are experimentally determined, the overall entropy generation is known. This provides the second needed equation, as the overall entropy generated must also equal the weighted sum of the family or elementary mode entropies.

$$\Delta S_{\text{total}} = \sum_j \Delta S_j\, c_j. \tag{9}$$

Substitution of (7) in (9) results in

$$\Delta S_{\text{total}} = a + b \sum_j c_j\, \ln\, c_j. \tag{10}$$

Which is an expression that permits computation of the total entropy on the basis of weights or probabilities of individual states. This result was somewhat surprising since this relationship is usually found in connection with the statistical thermodynamics of macroscopic systems comprised of a set of quantum states to relate the entropy of the system to the probabilities of quantum states.

This approach gave weighting factors for the PHB− system that are very close to the weighting factors computed from the mass balances. For the PHB+ set of families it suggests a balanced, entropy-based set of weighting factors that differ somewhat from the factors computed from the mass balances (Table 7).

To validate this approach further, we compared experimentally determined flux data obtained from literature with fluxes computed from elementary modes and weighting factors estimated from reaction entropies. This has been done for the wild type *E. coli* that operates under anaerobic non-growth conditions on the basis of 135 modes comprised of 35 families of modes, i.e. 35 weighting factors and associated elementary modes have been considered in the computation. The data show a clear correlation and are summarized in Fig. 7. This correlation appears to be highly significant considering that the experimental conditions between the sets of compared experimental data likely varied and probably contributed to the deviations from the correlation.

Each elementary mode operates according to a unique pathway with well-defined fluxes through individual reaction steps. To better visualize the degree of variation in individual fluxes in specific reaction steps specified in individual elementary modes, we have compared them to a set of fluxes determined experimentally (Fig. 8(a)). This figure shows there is a large degree of variation of fluxes between individual elementary modes, and these fluxes do not correlate with the experimentally determined rates. However, a weighted average of these fluxes correlates very nicely with the experimental data (Fig. 8(b)). It is emphasized again that the weights are determined only from the reaction entropies as described above. This result further validates the described approach to estimate metabolic fluxes from elementary modes.

5. Discussion

We have shown that it is possible to estimate fluxes through intracellular reactions on the basis of elementary

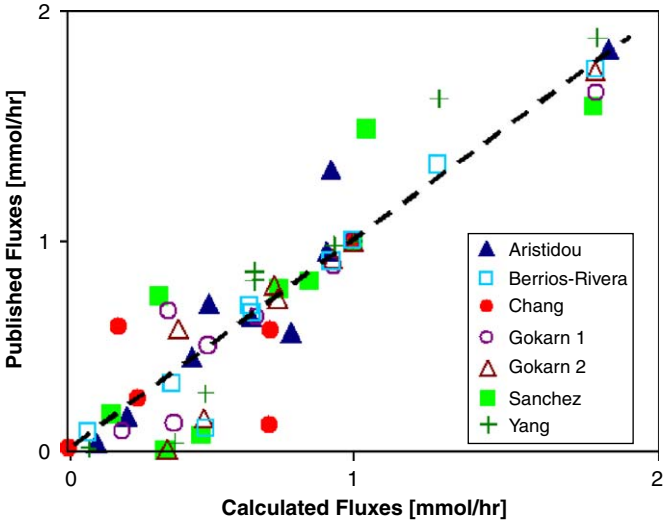

Fig. 7. Comparison between published and calculated fluxes. Legend refers to fluxes for wild type *E. coli* strains published in (Aristidou et al., 1999; Chang et al., 1999; Gokarn et al., 2000; Yang et al., 2001; Berrios-Rivera et al., 2002; Sanchez et al., 2005). Gokarn1 refers to the fluxes published for the VJS676 strain; Gokarn2 refers to the data published for the JCL1242/pPC201 strain (Gokarn et al., 2000). The published fluxes were normalized to the reported glucose uptake rate or flux through phosphofructokinease. The fluxes from Aristidou were extrapolated to a growth rate of zero from fluxes recorded from a series of chemostats at different growth rates. The Berrios-Rivera and Sanchez data corresponds to the overall fluxes with biosynthesis associated fluxes removed. Data from Chang comes from stationary phase experiments. The fluxes taken from Gorkarn and Yang correspond to the fluxes based on the carbon balance tied to the glucose consumption. The calculated fluxes are the fluxes through the adapted metabolically similar pathways represented by EMA. The overall correlation (dashed line) from all the data points has a slope of 0.99 and intercept of 6×10^{-6}. The coefficient of determination for this correlation is 0.85.

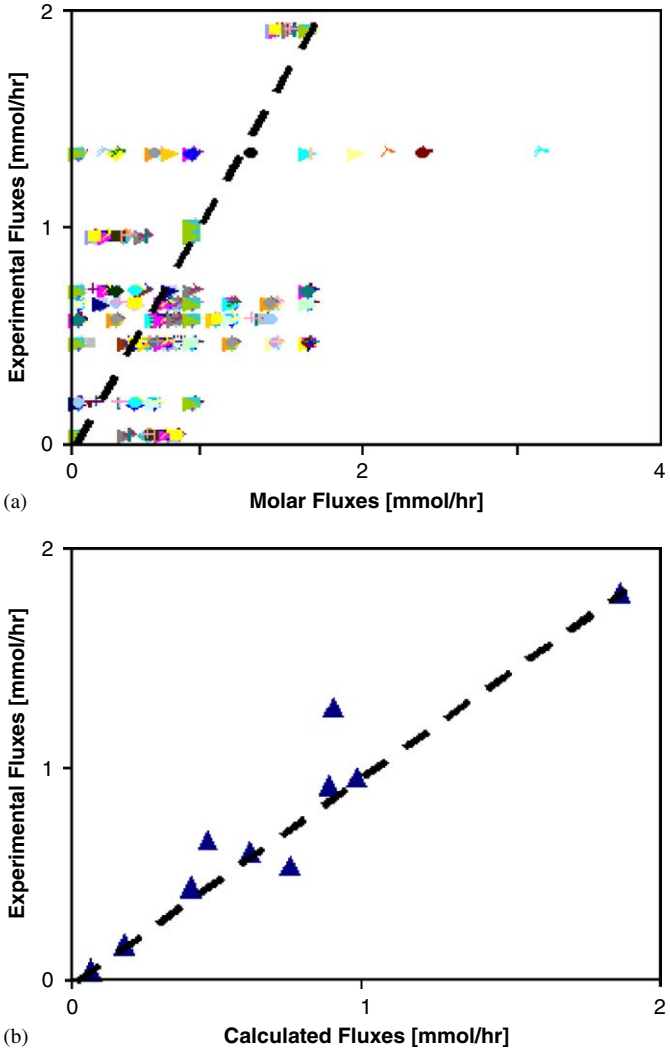

Fig. 8. Correlation between published fluxes from experiments (Aristidou et al., 1999) and calculated fluxes from EMA. (a) Experimentally determined fluxes versus fluxes per mole glucose consumed from individual elementary modes. The fluxes are normalized to the glucose uptake rate. (b) Correlation between experimentally measured fluxes (Aristidou et al., 1999) and fluxes computed as weighted averages of elementary modes. The weights have been estimated from reaction entropies (see text for details). The trend line in both figures is included for reference and has a slope of unity.

modes. This has been conventionally done using metabolic flux analysis, MFA. In MFA, experimentally determined secretion rates of specific metabolites are related to material balances specified by the reaction stoichiometry of the metabolic network. In special cases, a system defined in this way can be solved for intracellular fluxes. In contrast to this approach, we have applied a set of elementary modes to compute these intracellular fluxes.

Elementary modes define balanced fluxes through a metabolic network. Individual reaction fluxes in a cell represent weighted averages of all possible elementary modes present. To estimate the individual reaction rates it was necessary to estimate the unknown weighting factors associated with each elementary mode. This estimation is not possible in situations where the number of elementary modes outnumbers the number of metabolite accumulation rates that can be experimentally measured, unless there is additional information available. Therefore, we have implemented two gene deletions that reduce the number of possible elementary modes under the given growth conditions to 24 modes in the control strain and to 40 modes in the recombinant strain that is equipped with the PHB synthesis pathway. A further reduction of the number of unknown weighting factors has been achieved by the realization that certain elementary modes have the same overall stoichiometry and, therefore, can be grouped into a family of modes. For this system of reduced complexity the measured metabolites provided sufficient relationships so that the weighting factors could be calculated. The elementary modes within a family result in the same mass balance when the mode fluxes are related to the experimentally accessible secretion fluxes of external metabolites. Therefore, only a weighting factor for the entire family of modes can be determined from the measurement of secreted metabolites. It will be interesting to find out in future experiments how the individual elementary modes contribute to the overall fluxes of a

family. This information can be obtained through ^{13}C labeling studies or other in vivo techniques that can target specific reactions in the intracellular metabolism of a cell (Sauer et al., 1999).

In developing the analysis of our data we have indirectly derived relationships that commonly appear in statistical mechanics when a macroscopic system is comprised of discrete quantum states. The obtained results could suggest that the set of discrete elementary modes can be viewed as a complete set of possible states describing the metabolism of a cell. In statistical mechanics such set is known as a metabolic ensemble of states (Westerhoff and van Dam, 1987). For such states it is often possible to compute corresponding probabilities, which in turn permit computation of the entropies of individual states and the entropy of the system though a well-known logarithmic relationship that is analogous to the correlation obtained in this work. This correlation appears to determine a relationship according to which the overall entropy generation of the system is distributed among the discrete possible states.

It is remarkable that the reaction entropies of elementary modes correlate with the probabilities of their occurrence in this way. To make this happen, elementary modes must be subject to additional restrictions due to the dynamic nature of operational elementary modes. For an elementary mode to be in fact used, the catalysts of individual reaction steps must be present at the right proportions such that the overall flux through an elementary mode is supported. These properties are expected to be a result of billions of years of evolution that developed the underlying regulation and expression patterns of enzymes supporting specific reaction rates. A consequence of this thought is that it might be a basic principle of evolution to develop metabolic networks that enable a diversity of pathways but that favor pathways that generate the least amount of entropy. Some support of this hypothesis is the result that the PHB− cells (with Δ*frdA* Δ*ldh* genome containing only evolved pathways) show a perfect correlation between reaction entropies and elementary modes weights while the recombinant cells containing the transgenic PHB pathway presumably are not evolved and therefore the correlation is not that evident. This suggests that the tendency to reduce entropy generation within a metabolism still holds when pathways are removed. If these ideas were valid, they would have profound implications on the quantitative description of metabolic networks. It will remain to be shown in future studies whether these findings are indeed of more general nature or just artifacts of this specific experimental system.

Acknowledgments

The authors would like to thank Bernhard Sonnleitner for providing useful insight during the preparation of the manuscript. We are also grateful to the National Science Foundation (BES-0109383) and to the Minnesota Supercomputing Institute for supporting parts of this work. Ross Carlson and Aaron Wlaschin are recipients of NIH training grant fellowships in Biotechnology.

References

Aristidou, A.A., San, K.Y., Bennett, G.N., 1999. Metabolic flux analysis of *E. coli* expressing the bacillus subtilis acetolactate synthase in batch and continuous cultures. Biotechnol. Bioeng. 63, 737–749.

Berrios-Rivera, S.J., Bennett, G.N., San, K.Y., 2002. The effect of increasing NADH availability on the redistribution of metabolic fluxes in *E. coli* chemostat cultures. Metab. Eng. 4 (3), 230–237.

Blattner, F.R., Plunkett 3rd, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., Shao, Y., 1997. The complete genome sequence of *E. coli* K-12. Science 277 (5331), 1453–1474.

Burgard, A.P., Maranas, C.D., 2001. Probing the performance limits of the *E. coli* metabolic network subject to gene additions or deletions. Biotechnol. Bioeng. 74 (5), 364–375.

Carlson, R., Srienc, F., 2004a. Fundamental *E. coli* biochemical pathways for biomass and energy production: creation of overall flux states. Biotechnol. Bioeng. 86 (2), 149–162.

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

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

Carlson, R., Wlaschin, A., Srienc, F., 2005. Kinetic studies and biochemical pathway analysis of anaerobic poly (R)-3-hydroxybutyric acid synthesis in *E. coli*. Appl. Environ. Microbiol. 71 (2), 713–720.

Chang, D.E., Jung, H.C., Rhee, J.S., Pan, J.G., 1999. Homofermentative production of D- or L-lactate in metabolically engineered *E. coli* Rr1. Appl. Environ. Microbiol. 65 (4), 1384–1389.

Choi, J.I., Lee, S.Y., Han, K., 1998. Cloning of the *Alcaligenes latus* polyhydroxyalkanoate biosynthesis genes and use of these genes for enhanced production of poly(3-hydroxybutyrate) in *E. coli*. Appl. Environ. Microbiol. 64 (12), 4897–4903.

Clark, D.P., 1989. The fermentation pathways of *E. coli*. FEMS Microbiol. Rev. 5 (3), 223–234.

Dandekar, T., Schuster, S., Snel, B., Huynen, M., Bork, P., 1999. Pathway alignment: application to the comparitive analysis of glycolytic enzymes. J. Biochem. 343, 115–124.

Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *E. coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA 97 (12), 6640–6645.

Edwards, J.S., Palsson, B.O., 2000. Robustness analysis of the *E. coli* metabolic network. Biotechnol. Prog. 16 (6), 927–939.

Eschenlauer, A.C., Stoup, S.K., Srienc, F., Somers, D.A., 1996. Production of heteropolymeric polyhydroxyalkanoate in *E. coli* from a single carbon source. Int. J. Biol. Macromol. 19 (2), 121–130.

Gokarn, R.R., Eiteman, M.A., Altman, E., 2000. Metabolic analysis of *E. coli* in the presence and absence of the carboxylating enzymes phosphoenolpyruvate carboxylase and pyruvate carboxylase. Appl. Environ. Microbiol 66 (5), 1844–1850.

Heinrich, R., Schuster, S., 1996. The Regulation of Cellular Systems. Chapman & Hall, Cincinnati.

Heldal, M., Norland, S., Tumyr, O., 1985. X-ray microanalytic method for measurement of dry matter and elemental content of individual bacteria. Appl. Environ. Microbiol. 50 (5), 1251–1257.

Jackson, D.E., Srienc, F., 1994. Novel methods to synthesize polyhydroxyalkanoates. Ann. NY. Acad. Sci. 745, 134–148.

Kawasaki, T., Nakata, T., Nose, Y., 1968. Genetic mapping with a thiamine-requiring auxotroph of *E. coli* K-12 defective in thiamine phosphate pyrophosphorylase. J. Bacteriol. 95 (4), 1483–1485.

Kidwell, J., Valentin, H.E., Dennis, D., 1995. Regulated expression of the *Alcaligenes eutrophus* PHA biosynthesis genes in *E. coli*. Appl. Environ. Microbiol. 61 (4), 1391–1398.

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

Leaf, T.A., Srienc, F., 1998. Metabolic modeling of polyhydroxybutyrate biosynthesis. Biotechnol. Bioeng. 57 (5), 557–570.

Lee, S.Y., Choi, J.I., 2001. Production of microbial polyester by fermentation of recombinant microorganisms. Adv. Biochem. Eng. Biotechnol. 71, 183–207.

Lee, S.Y., Papoutsakis, E.T. (Eds.), 1999. Metabolic Engineering. Marcel Dekker, Inc., New York.

Lee, S.Y., Lee, Y.K., Chang, H.N., 1995. Stimulatory effects of amino acids and oleic acid on poly(3-hydroxybutyric acid) synthesis by recombinant *E. coli*. J. Ferment. Bioeng. 79 (2), 177–180.

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

Mainzer, S.E., Hempfling, W.P., 1976. Effects of growth temperature on yield and maintenance during glucose-limited continuous culture of *E. coli*. J. Bacteriol. 126 (1), 251–256.

Maklashina, E., Berthold, D.A., Cecchini, G., 1998. Anaerobic expression of *E. coli* succinate dehydrogenase: functional replacement of fumarate reductase in the respiratory chain during anaerobic growth. J. Bacteriol. 180 (22), 5989–5996.

Mori, H., Isono, K., Horiuchi, T., Miki, T., 2000. Functional genomics of *E. coli* in Japan. Res. Microbiol. 151 (2), 121–128.

Neidhardt, F.C. (Ed.), 1996. *E. coli* and *Salmonella*: Cellular and Molecular Biology. ASM Press, Washington, DC.

Papin, J.A., Stelling, J., Price, N.D., Klamt, S., Schuster, S., Palsson, B.O., 2004. Comparison of network-based pathway analysis methods. Trends Biotechnol. 22 (8), 400–405.

Peoples, O.P., Sinskey, A.J., 1989. Poly beta-hydroxybutyrate (PHB) biosynthesis in *Alcaligenes eutrophus* H16. identification and characterization of the PHB polymerase gene (phbc). J. Biol. Chem. 264 (26), 15298–15303.

Pfeiffer, T., Sanchez-Valdenebro, I., Nuno, J.C., Montero, F., Schuster, S., 1999. Metatool: for studying metabolic networks. Bioinformatics 15 (3), 251–257.

Pirt, S.J., 1975. Principles of Microbe and Cell Cultivation. Wiley, New York.

Poolman, M.G., Venkatesh, K.V., Pidcock, M.K., Fell, D.A., 2004. A method for the determination of flux in elementary modes, and its application to *Lactobacillus rhamnosus*. Biotechnol. Bioeng. 88 (5), 601–612.

Pramanik, J., Keasling, J.D., 1998. Effect of *E. coli* biomass composition on central metabolic fluxes predicted by a stoichiometric model. Biotechnol. Bioeng. 60 (2), 230–238.

Riis, V., Mai, W., 1988. Gas chromatographic determination of poly B-hydroxybutyric acid in microbial biomass after hydrochloric acid propanolysis. J. Chromatogr. 445, 285–289.

Roels, J.A., 1983. Energetics and Kinetics in Biotechnology. Elsevier Biomedical Press, New York.

Sanchez, A.M., Bennett, G.N., San, K.Y., 2005. Effect of different levels of NADH availability on metabolic fluxes of *E. coli* chemostat cultures in defined medium. J. Biotech. 117 (4), 395–405.

Sandler, S., Orbey, H., 1991. On the thermodynamics of microbial growth processes. Biotechnol. Bioeng. 38 (7), 697–718.

Sauer, U., Lasko, D.R., Fiaux, J., Hochuli, M., Glaser, R., Szyperski, T., Wuthrich, K., Bailey, J.E., 1999. Metabolic flux ratio analysis of genetic and environmental modulations of *E. coli* central carbon metabolism. J. Bacteriol. 181 (21), 6679–6688.

Schubert, P., Steinbuchel, A., Schlegel, H.G., 1988. Cloning of the *Alcaligenes eutrophus* genes for synthesis of poly beta-hydroxybutyric acid (PHB) and synthesis of PHB in *E. coli*. J. Bacteriol. 170 (12), 5837–5847.

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

Schuster, S., Fell, D.A., 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.

Sipkema, E.M., de Koning, W., Ganzeveld, K.J., Janssen, D.B., Beenackers, A.A., 2000. NADH-regulated metabolic model for growth of *Methylosinus trichosporium* Ob3b. Model presentation, parameter estimation, and model validation. Biotechnol. Prog. 16 (2), 176–188.

Steinbuchel, A., 1991. Biomaterials: Novel Materials from Biological Sources. Stockton Press, New York.

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

Sudesh, K., Abe, H., Doi, Y., 2000. Synthesis, structure and properties of polyhydroxyalkanoates: biological polyesters. Prog. Polym. Sci. 25, 1503–1555.

Talu, G.F., Diyamandoglu, V., 2004. Formate ion decomposition in water under UV irradiation at 253.7 nm. Environ. Sci. Technol. 38 (14), 3984–3993.

Trinh, C., Carlson, R., Wlaschin, A., Srienc, F., 2006. Design, construction and performance of the most efficient *E. coli* bacterium, submitted.

van Wegen, R.J., Ling, Y., Middelberg, A.P., 1998. Industrial production of polyhydroxyalkanoates using *E. coli*: an economic analysis. Trans. Chem. Eng. 76, 417–426.

van Wegen, R.J., Lee, S.Y., Middelberg, A.P., 2001. Metabolic and kinetic analysis of poly(3-hydroxybutyrate) production by recombinant *E. coli*. Biotechnol. Bioeng. 74 (1), 70–80.

Vijayasankaran, N., Carlson, R., Srienc, F., 2005. Metabolic pathway structures for recombinant protein synthesis in *E. coli*. Appl. Microbiol. Biotechnol. 68, 737–746.

Vollbrecht, D., El Nawawy, M.A., Schlegel, H.G., 1978. Excretion of metabolites by hydrogen bacteria: I autotrophic and heterotrophic fermentations. Eur. J. Appl. Microbiol. 6, 145–155.

Wang, F., Lee, S.Y., 1998. High cell density culture of metabolically engineered *E. coli* for the production of poly(3-hydroxybutyrate) in a defined medium. Biotechnol. Bioeng. 58 (2–3), 325–328.

Westerhoff, H.V., van Dam, K., 1987. Thermodynamics and Control of Biological Free-Energy Transduction. Elsevier Science Publication, New York.

Yang, Y.T., Bennett, G.N., San, K.Y., 2001. The effects of feed and intracellular pyruvate levels on the redistribution of metabolic fluxes in *E. coli*. Metab. Eng. 3 (2), 115–123.