MFA for Overdetermined Systems Reviewed and Compared with RNA Expression Data to Elucidate the Difference in Shikimate Yield between Carbon- and Phosphate-Limited Continuous Cultures of *E. coli* W3110.shik1

Gaspard Lequeux,*,† Louise Johansson,‡ Jo Maertens,† Peter A. Vanrolleghem,† and Gunnar Lidén‡

BIOMATH, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure Links 653, B-9000 Gent, Belgium, and Department of Chemical Engineering, Lund University, P.O. Box 124, SE-22100 Lund, Sweden

The present contribution focuses on the mathematical techniques used to solve steady state metabolic models for the case of an overdetermined system. Even when parts of the system are underdetermined it is possible to solve the model partially and obtain statistically meaningful results. This is illustrated with data gathered from a set of *E. coli* W3110.shik1 phosphate- or carbon-limited continuous cultures. It is shown that the low yield in shikimate for C-limited cultures is not due to a lower flux going to the shikimate pathway but is caused by a high secretion of byproducts. Carbon-limited cultures could be better for shikimate production than carbon-abundant cultures provided the byproduct secretion is reduced. Finally, flux calculations are compared with RNA expression data.

1. Introduction

Shikimic acid is an interesting starting material for the production of many chemical compounds, with as major example Tamiflu, used for treatment of influenza (*1*). As the extraction of shikimate from the plant *Illicium* is expensive, production strategies via fermentation are developed. One of the strains that has been genetically modified to produce more shikimate is *E. coli* W3110.shik1 (*2*). Shikimate production, however, is very low under carbon-limiting conditions.

Metabolic flux analysis can be used to obtain more knowledge about the flux distribution of intracellular reactions. MFA is frequently applied on underdetermined systems when the solution of the model can only be calculated under certain assumptions, e.g., maximal biomass production (*3*). However, metabolic models can also be used when the system is overdetermined, i.e., when there are more measurements available than degrees of freedom (*4*). The statistical methods discussed in this paper were initially developed for black box models (*5*, *6*) but can easily be applied on metabolic models (*7*).

In the first part of this contribution, those methods are reviewed. It will be clarified that one can still calculate some fluxes and perform statistical tests to check whether the measurement data are consistent even if parts of the model are underdetermined (*7*). The methods explained here are more general than what can be found in refs 8 and 9.

In a second part, those methodologies are applied to a set of experiments aiming at finding differences in flux distribution between carbon-limited and carbon-abundant (phosphate-limited) cultures of the *E. coli* strain W3110.shik1 (*2*). This strain was genetically modified to produce more shikimate (an intermediate of the aromatic amino acid pathway). Under carbon-limited conditions, more byproducts can be found, such as dehydro shikimate and dehydro quinate, and a lot less shikimate is produced (*2*, *10*) than under carbon-abundance. It is shown that this lower shikimate yield under carbon-limited conditions is not due to a lower flux going to the aromatic pathway. On the contrary there is an even higher flux going to it.

The last part of this work compares flux balancing with RNA expression level measurements. Although it is sometimes attempted to solve the problem of parallel pathways with the aid of RNA expression levels (*11*), this is not done here because it is not expected to be possible (*12*). However, for some fluxes the variations over different culture conditions correlates with the RNA expression levels of the corresponding genes.

2. Materials and Methods

2.1. Experimental Setup. In the performed experiments the *E. coli* W3110 strain was used. This strain was modified to W3110.shik1, as described in ref 2.

Six phosphate-limited and four carbon-limited continuous cultures with W3110.shik1 were run at dilution rates varying from 0.05 to 0.3 h^{-1}.

A second set of experiments consisted of phosphate- and carbon-limited chemostats of the wild-type W3110 and the shikimate-producing strain W3110.shik1 at a dilution rate of 0.2 h^{-1}. Each combination was done in duplicate. Besides input and output fluxes, RNA expression levels were determined too.

2.1.1. ***Inoculum Preparation.*** The medium used for inoculum preparation (100 mL) was the same for carbon-limited and phosphate-limited experiments. For W3110 (the wild-type strain), the medium contained, per liter: 20 g of glucose (VWR International, Stockholm, Sweden), 7.5 g of K_2HPO_4, 0.3 g of NH_4Fe(III)-citrate, 2.1 g of citric acid monohydrate, 0.1 mL $MgSO_4 \cdot 7\ H_2O$ (1 M), 0.0037 g of $(NH_4)_6(Mo_7O_{24}) \cdot 4\ H_2O$, 0.0029 g of $ZnSO_4 \cdot 7\ H_2O$, 0.0247 g of H_3BO_3, 0.0025 g of $CuSO_4 \cdot 5\ H_2O$, 0.0158 g of $MnCl_2 \cdot 4\ H_2O$, 0.055 mg of thiamine (Sigma-Aldrich, Steinheim, Germany), and 1.2 mL of H_2SO_4 (96%) (VWR International). For W3110.shik1 the medium additionally contained 0.010 g of *p*-hydroxybenzoic acid, 0.010

* To whom correspondence should be addressed. gaspard.lequeux@biomath.ugent.be.
† Ghent University.
‡ Lund University.

10.1021/bp060014u CCC: $33.50
Published on Web 06/16/2006

g of potassium *p*-amino benzoate, 0.010 g of 2,3-dihydroxybenzoic acid (vitamins), and 0.015 g of tetracycline hydrochloride (antibiotic) (Sigma-Aldrich). Inoculum cultures were grown in 250 mL baffled E-flasks at 37 °C at a shaker speed of 200 rpm up to an OD-value of 1.5−2.5, to ensure exponentially growth. The fermenter was inoculated with 40 mL of the inoculum culture (W3110.shik1) or 20 mL (W3110).

2.1.2. *Culture Media.* The glucose and mineral solutions were sterilized separately at 121 °C for 20 min and thereafter mixed. Antibiotic solution, vitamins and trace metals, and $MgSO_4$ solution were added by sterile filtration through a 0.2 mm Minisart cellulose acetate filter (Sartorius AG, Goettingen, Germany). All solutions were prepared using deionized water; 1.5 L of medium was prepared for the batch cultivation. The working volume of the chemostat was around 1.3 L, and pH was adjusted to 7.0 by addition of NH_4OH (25%) before and after sterilization.

2.1.3. *Carbon-Limited Cultivations.* The composition of the C-limitation medium for the batch phase and the chemostat phase of the W3110 was the same as for the preculture medium, except for some minor changes. The medium contained 0.13 g/L antifoam 286 (Sigma-Aldrich). In addition, the glucose concentration was 10 and 25 g/L in batch and chemostat phase, respectively. The composition of the medium for growth of W3110.shik1 was the same as that for W3110 except that it also contained 0.015 g/L tetracycline hydrochloride.

2.1.4. *Phosphate-Limited Cultivations.* The composition of the P-limited medium in batch phase was, per liter: 20 g of glucose, 92 mL of H_3PO_4 (85%), 5.39 g of NH_4SO_4, 3.32 g of NaOH, 1.66 g of KOH (Sigma-Aldrich), 0.52 g of $MgSO_4 \cdot 7\ H_2O$, 0.133 g of antifoam 286, 3.73 mL of H_2SO_4 (96%); trace metals (1): 0.093 g of $FeSO_4 \cdot 7\ H_2O$, and 0.079 g of citric acid monohydrate; trace metal (2): 0.0073 g of $CoCl_2 \cdot 6\ H_2O$, 0.00207 g of $MnCl_2 \cdot 4\ H_2O$, and 0.00103 g of ZnCl (Sigma-Aldrich). The medium of W3110.shik1 also contained 0.015 g/L tetracycline hydrochloride. In chemostat operation, the phosphate solution and the solution of the remaining medium components were fed separately to the fermenter. The phosphate feed contained 0.9 mL/L H_3PO_4 (85%), whereas the second feed solution contained all other components. Since the phosphate solution contributed extra volume and thereby diluted the second feed solution the latter was up concentrated 1.3 times, e.g., the glucose concentration was 32 g/L, giving a concentration of 25 g/L in the reactor.

2.1.5. *Fermentation Methods.* All fermentations were carried out in a Biostat CT culture vessel (Sartorius BBI Systems GmbH, Melsungen, Germany) with a maximum working volume of 3.5 L. Temperature (37 °C), pH (7.0), stirring rate (750 rpm), and airflow rate (0.75 slpm) were controlled by the program MFCS/win shell 2.0 (Sartorius BBI) via the control unit DCU Biostat C (Sartorius BBI). p O_2 was measured with an InPro 6000 (Mettler Toledo Gmbh, Giessen, Germany). The pH was measured with a pH-meter of type 405-DPAS-SC-K8S/120 (Mettler Toledo). For maintaining the pH at 7.0, H_2SO_4 (2 M) and NH_4OH (25%) were used in the C-limited case and H_2SO_4 (1 M) and NH_4OH (12.5%) in the P-limited case. O_2 and CO_2 contents in the off gas were measured by an Innova 1311 (INNOVA Air Tech Instruments, Ballerup, Denmark). Two balances and pumps were also coupled to the equipment, allowing precise measurement of the feed rate during chemostat operation. Data from all equipment except for one of the pumps were routinely logged by the data acquisition software.

In the C-limited case, chemostat operation was started when the glucose was finished (indicated by a rapid reduction of the CO_2 in the off-gas). In the P-limited case, the chemostat was started when the phosphate was consumed (also seen as a small peak in the p O_2). Establishment of steady state was confirmed from measurement of the off-gas composition. At least five retention times were allowed to pass before each steady-state. Sampling of OD and for metabolite analyses was carried out at fermentation start, end of batch, and at steady state during chemostat operation. Samples for dry weight were taken at the end of the batch cultivation and at steady state during chemostat cultivation.

2.2. Analytical Methods. The analytical methods for determining the amount of metabolites and cell dry weights are described in ref 2; 3.5% of the dry weight was considered as ash (*13*).

For the elemental composition of biomass, 50 mL of fermentation broth was taken and centrifuged for 20 min at 2000 *g* and 4 °C. The pellet was washed once with 40 mL of ice cold 0.9% NaCl and again centrifuged and decanted. The pellets were then frozen in liquid nitrogen, vacuum-dried, and analyzed with a CHNS-O analyzer (model EA1108, Carlo Erba Instruments, Italy) (*4*).

Sampling for transcriptome analysis was performed at steady state. To minimize mRNA degradation, 50 mL falcon tubes containing about 25 mL of ice were filled up as fast as possible with fermentation broth. The tubes were centrifuged for 1 min at 2000 *g*, and the pellet was frozen in liquid nitrogen. In the phosphate-limited experiments, the mRNA sampling method was further improved by adding 1.25 mL of "stop-solution", containing about 95% ethanol and 5% phenol, to the tubes prior to sampling.

Total RNA was extracted by using the Fastprep system including the FastRNAPro Blue kit (Qbiogene, Montréal, Canada). DNA present was degraded by addition of DNAse (VWR International) to a concentration of 0.2 U μL^{-1} . The samples were then held at 37 °C for 20 min. The reaction was interrupted by addition of 0.5 M EDTA (Sigma-Aldrich) to a final concentration of 10 mM. The RNA was further cleaned by using a RNeasy microelute Cleanup kit (Qiagen, Venlo, The Netherlands). The quality of RNA was controlled by running the samples on a 1% TBE gel. The RNA samples were then sent to SWEGENE Microarray Resource Centre (Lund, Sweden) where the quality of RNA was further controlled by using Nanodrop ND1000 (Nanodrop Technologies, Wilmington, USA), which gives very accurate concentration and 260/280 ratio figures. In addition quality testing of the RNA was carried out using the Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, USA). cDNA synthesis, cDNA fragmentation, and preparation of the hybridization mixture was carried out according to the recommendations of the manufacturer of the microarrays (Affymetrix, Santa Clara, USA). Hybridization, washing, staining, and scanning of the microarrays (Affymetrix *E. coli* antisense genome arrays AS v2) were performed by using GeneChip Hybridization Oven 640, GeneChip Fluidic Station 450, and GeneChip Scanner 2500 (all from Affymetrix). Data acquisition and gene expression data analysis was carried out using MAS 5.0 (Affymetrix). A chip to chip normalization was performed in this program by scaling to a median intensity of 100. For statistical analysis the Bayesian test was carried out using a web-interface version of Cyber-T found at http://visitor.ics.uci.edu/genex/cybert/ (*14*). All presented microarray data were significant at a 95% level. The mRNA data are discussed in more detail in ref 15.

2.3. Theoretical Framework. A metabolic network model is basically a system of linear equations. There are two methods

to solve such a system. The first one works directly with nullspace calculations (*16*). It is an elegant method but has a major drawback: it is not possible to use internal flux measurements to reduce the number of degrees of freedom in the system. Only exchange rates can be used. The second method (well described in ref 17) is more general and is the one that is used in this work. Both methods make use of the pseudo inverse of a matrix and its nullspace. The methods will be reviewed briefly.

2.3.1. ***Pseudo Inverse of a Matrix.*** The solution of a linear system

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b} \tag{1}$$

can be found by calculating the inverse of matrix **A**:

$$\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b} \tag{2}$$

This can only be done if **A** is invertible, which means that the system of equations should be neither overdetermined nor underdetermined. In MFA those conditions are rarely fulfilled.

In the case of an overdetermined system a "least-squares" solution can be calculated by use of the (left) pseudo inverse:

$$\boldsymbol{x} = (\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\,\boldsymbol{b} \tag{3}$$

The left pseudo inverse can thus be defined as $(\mathbf{A}^{\mathbf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}$ and to be calculable, $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ must be nonsingular. This means that the number of rows of **A** should be less or equal to the number of columns. Furthermore, those rows should be linearly independent.

A more general method, which can be applied to every matrix, to calculate the pseudo inverse is based on Singular Value Decomposition (*18*) (a good explanation of SVD, the pseudo inverse, and its application in MFA can be found in ref 19). Given the matrix **A** with dimensions $\langle m \times n \rangle$, there exist orthonormal matrixes:

$$\mathbf{U}_{\langle m\times m\rangle} \text{ and } \mathbf{V}_{\langle n\times n\rangle} \tag{4}$$

so that

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}} \tag{5}$$

where **S** is an $\langle m \times n \rangle$ dimensional nonsquare diagonal matrix. The elements on the diagonal are the singular values of **A**. The number of singular values that are not zero is equal to the rank of **A**. When **A** has full rank and $m = n$, the inverse of **A** is (the inverse of an orthonormal matrix is equal to its transposed)

$$\mathbf{A}^{-1} = \mathbf{V}\mathbf{S}^{-1}\mathbf{U}^{\mathbf{T}} \tag{6}$$

The inverse of a diagonal matrix is calculated by inverting the elements on the diagonal one by one. If **A** is not fully ranked or **A** is not square, a partial inverse can be calculated by inverting only singular values in **S** that are not zero. This way the pseudo inverse of every matrix can be calculated:

$$\mathbf{A}^{\#} = \mathbf{V}\mathbf{S}^{\#}\mathbf{U}^{\mathbf{T}} \tag{7}$$

This makes it possible to calculate a solution not only when the system is fully or overdetermined but also when parts of the system are underdetermined. The next section explains how one can make a distinction, by use of the nullspace, between those underdetermined parts (whose obtained solution is obviously not unique) and the determined ones.

2.3.2 ***Nullspace.*** If the system as given in eq 1 is underdetermined, the solution given with the pseudo inverse

$$\boldsymbol{x} = \mathbf{A}^{\#}\boldsymbol{b} \tag{8}$$

is only one of the infinity of possible solutions. To determine which elements of $\boldsymbol{x}$ have a single solution and to determine a relationship between the infinity of solutions for the other elements of $\boldsymbol{x}$, the nullspace can be calculated.

The nullspace is defined as the set of linear independent vectors, $\boldsymbol{x}_n$, that fulfill the equation

$$\mathbf{A}\boldsymbol{x}_n = 0 \tag{9}$$

The number of independent nullspace vectors is equal to the number of columns in **A** minus the rank of **A**. From the definition it is clear that each nullspace vector can be added an arbitrary number of times to the base solution given in eq 8. Thus the complete solution is

$$\boldsymbol{x} = \mathbf{A}^{\#}\boldsymbol{b} + \text{nullspace}(\mathbf{A})\boldsymbol{f} \tag{10}$$

where $\boldsymbol{f}$ is a vector with as many elements as there are vectors (columns) in the nullspace of **A**. For all possible values of $\boldsymbol{f}$ the solution remains valid.

Parallel pathways in the metabolic model typically yield nullspace vectors if there are no fluxes measured from one of those pathways. Elements (rows) of $\boldsymbol{x}$ that have only zeros in the nullspace of **A** are fully determined. Thus, a system of equations can have some unknowns for which a unique solution can be found and some for which this is not possible.

2.3.3. ***Construction of the Metabolic Model.*** When gathering the reactions for the metabolic model, inevitably some errors get in. Errors in stoichiometry were caught with the elemental consistency test. Superfluous reactions were detected with the dead end test and removed. Both tests are described elsewhere (*20*).

Although the technique for solving the metabolic model explained further can cope with parallel pathways, we chose to remove them, to be able to fully solve the model and have all fluxes known. During the construction of the model, parallel pathways were detected with the method described in ref 20 and removed.

2.3.4. ***Solution of the Metabolic Model.*** The matrix representation of a biochemical reaction network can be written as

$$\frac{\mathrm{d}\boldsymbol{c}}{\mathrm{d}t} = \mathbf{W}\boldsymbol{a}(\boldsymbol{c}) \tag{11}$$

were $\boldsymbol{c}$ is a vector containing the concentration of the different metabolites inside the cell, matrix **W** is the extended stoichiometric matrix, and vector $\boldsymbol{a}$ is the vector with all the reaction rates. If steady state is assumed, the concentration of the metabolites inside the cell does not change and eq 11 can be simplified to

$$\mathbf{W}\boldsymbol{a} = 0 \tag{12}$$

W is called the extended stoichiometric matrix because it contains more than pure enzymatic conversions. The exchange of metabolites with the environment is added as reactions taking from, or removing matter to a buffer (*21*). Thus input and output

from the system is modeled as (e.g. glucose):

$$\rightarrow \mathrm{GLC} \tag{13}$$

The corresponding rate in the flux vector $\boldsymbol{a}$ is the input rate for glucose.

In MFA the following representation is frequently used for a metabolic model (*8*, *9*):

$$\begin{bmatrix} \boldsymbol{r} \\ \mathbf{0} \end{bmatrix} = \mathbf{S}\boldsymbol{v} \tag{14}$$

where $\boldsymbol{r}$ is the net accumulation rate vector, $\boldsymbol{v}$ is the reaction rate vector, sensu stricto, and $\mathbf{S}$ is the stoichiometric matrix with the different reactions in the rows and the metabolites in the columns. This equation can be rewritten as ($\mathbf{I}$ representing the identity matrix augmented with some rows containing zeros, corresponding to the metabolites that have a zero exchange rate in eq 14):

$$[\mathbf{S} - \mathbf{I}] \begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{r} \end{bmatrix} = 0 \tag{15}$$

which matches eq 12.

Equation 12 is split up in two parts: one for the rates that are measured and thus known (subscript m) and one for the rates to be calculated (subscript c):

$$\mathbf{W}_\mathrm{c}\boldsymbol{a}_\mathrm{c} + \mathbf{W}_\mathrm{m}\boldsymbol{a}_\mathrm{m} = 0 \tag{16}$$

This equation can be solved for $\boldsymbol{a}_\mathrm{c}$:

$$\boldsymbol{a}_\mathrm{c} = -\mathbf{W}_\mathrm{c}^{\#}\mathbf{W}_\mathrm{m}\boldsymbol{a}_\mathrm{m} + \mathrm{nullspace}(\mathbf{W}_\mathrm{c})\boldsymbol{f} \tag{17}$$

2.3.5. ***Redundant Measurements.*** The solution obtained in eq 17 can be substituted into eq 16:

$$\mathbf{W}_\mathrm{m}\boldsymbol{a}_\mathrm{m} - \mathbf{W}_\mathrm{c}(\mathbf{W}_\mathrm{c}^{\#}\mathbf{W}_\mathrm{m}\boldsymbol{a}_\mathrm{m} + \mathrm{nullspace}(\mathbf{W}_\mathrm{c})\boldsymbol{f}) = 0 \tag{18}$$

As a matrix multiplied with its nullspace is equal to zero, this equation can be rewritten as

$$(\mathbf{W}_\mathrm{m} - \mathbf{W}_\mathrm{c}\mathbf{W}_\mathrm{c}^{\#}\mathbf{W}_\mathrm{m})\boldsymbol{a}_\mathrm{m} = 0 \tag{19}$$

When the system is overdetermined, i.e., when there are more measurements than degrees of freedom (and assuming the measurements do not agree perfectly), the terms before $\boldsymbol{a}_\mathrm{m}$ in eq 19 form a nonzero matrix. All independent rows of that matrix can be combined in the redundancy matrix $\mathbf{R}$ (*6*, *22*):

$$\mathbf{R}\boldsymbol{a}_\mathrm{m} = 0 \tag{20}$$

It should be noted that it is not because there is redundancy in the measurements that the nullspace of the extended stoichiometric matrix should be void. It is perfectly possible (and usual) to have parts of the system of equations that are overdetermined and parts that are underdetermined.

2.3.6. ***Making Use of Redundant Measurements.*** Redundant measurements can be used to enhance the confidence in the measurements. This is explained in ref 5 for black box models, but the method also applies to stoichiometric models (*7*). The real value of a measurement $\boldsymbol{a}_\mathrm{m}$ is equal to the measured value $\tilde{\boldsymbol{a}}_\mathrm{m}$ minus some random noise $\boldsymbol{\delta}$:

$$\boldsymbol{a}_\mathrm{m} = \tilde{\boldsymbol{a}}_\mathrm{m} - \boldsymbol{\delta} \tag{21}$$

Equation 20 gives the general formula; applied in a statistical context, distinction should be made between a real value of a variable (that can never be known), the measurement of that variable (indicated with a tilde), and the estimate of that variable (indicated with a hat). Combining eqs 21 and 20 yields

$$\mathbf{R}\,\tilde{\boldsymbol{a}}_\mathrm{m} = \mathbf{R}(\boldsymbol{a}_\mathrm{m} + \boldsymbol{\delta}) = \mathbf{R}\boldsymbol{\delta} = \boldsymbol{\epsilon} \tag{22}$$

$\boldsymbol{\epsilon}$ being the vector of residuals. It was proven that to minimize the error $\boldsymbol{\delta}$ on the measurements, the following objective function J has to be minimized (*23*):

$$J = \boldsymbol{\delta}^\mathbf{T}\mathbf{P}_\epsilon^{-1}\boldsymbol{\delta} \tag{23}$$

$\mathbf{P}_\epsilon$ is the variance covariance matrix of the vector of residuals, and it can be proven that

$$\mathbf{P}_\epsilon = \mathbf{R}\mathbf{P}_\delta\mathbf{R}^\mathbf{T} \tag{24}$$

The variance covariance matrix of $\boldsymbol{\delta}$ ($\mathbf{P}_\delta$) is equal to the variance covariance matrix of the measurements $\mathbf{P}_{\tilde{a}_\mathrm{m}}$.

The solution of the minimization problem of eq 23 gives an estimate for $\boldsymbol{\delta}$:

$$\hat{\boldsymbol{\delta}} = \mathbf{P}_\delta\,\mathbf{R}^\mathbf{T}\,\mathbf{P}_\epsilon^{-1}\,\boldsymbol{\epsilon} \tag{25}$$

which in turn gives an estimate for the measurements ($\mathbf{I}$ is the identity matrix):

$$\hat{\boldsymbol{a}}_\mathrm{m} = (\mathbf{I} - \mathbf{P}_{\tilde{a}_\mathrm{m}}\mathbf{R}^\mathbf{T}(\mathbf{R}\mathbf{P}_{\tilde{a}_\mathrm{m}}\mathbf{R}^\mathbf{T})^{-1}\mathbf{R})\tilde{\boldsymbol{a}}_\mathrm{m} \tag{26}$$

That these estimated values are better than the measured ones is obvious from the variance covariance matrix of the estimated values:

$$\mathbf{P}_{\hat{a}_\mathrm{m}} = (\mathbf{I} - \mathbf{P}_{\tilde{a}_\mathrm{m}}\mathbf{R}^\mathbf{T}(\mathbf{R}\mathbf{P}_{\tilde{a}_\mathrm{m}}\mathbf{R}^\mathbf{T})^{-1}\mathbf{R})\mathbf{P}_{\tilde{a}_\mathrm{m}} \tag{27}$$

since the second term of $\mathbf{P}_{\hat{a}_\mathrm{m}}$ is always positive. Making an estimate of the measured fluxes thus reduces their uncertainty.

The estimated measurements should be used to calculate the unknown fluxes. Equation 17 becomes

$$\hat{\boldsymbol{a}}_\mathrm{c} = -\mathbf{W}_\mathrm{c}^{\#}\mathbf{W}_\mathrm{m}\,\hat{\boldsymbol{a}}_\mathrm{m} + \mathrm{nullspace}(\mathbf{W}_\mathrm{c})\boldsymbol{f} \tag{28}$$

and the variance covariance matrix of the unknown fluxes becomes (*19*)

$$\mathbf{P}_{\hat{a}_\mathrm{c}} = \mathbf{W}_\mathrm{c}^{\#}\mathbf{W}_\mathrm{m}\mathbf{P}_{\hat{a}_\mathrm{m}}\mathbf{W}_\mathrm{m}^\mathbf{T}\mathbf{W}_\mathrm{c}^{\#\mathbf{T}} \tag{29}$$

Only the rows and columns of $\mathbf{P}_{\hat{a}_\mathrm{c}}$ corresponding to elements of $\hat{\boldsymbol{a}}_\mathrm{c}$ that have no freedom left in the nullspace of $\mathbf{W}_\mathrm{c}$ (i.e., for which the corresponding rows in the nullspace of $\mathbf{W}_\mathrm{c}$ contain only zeros) are relevant.

2.3.7. ***Statistical Test of the Quality of the Measurements.*** In the previous section it was explained how to increase the reliability of the measured fluxes. It is also possible to check whether the measurements are consistent. If so the residual vector ϵ should be equal to zero. To investigate whether this is the case, the following test statistic h is used (*24*):

$$h = \boldsymbol{\epsilon}^\mathbf{T}\mathbf{P}_\epsilon^{-1}\boldsymbol{\epsilon} \tag{30}$$

The H_0 hypothesis of this statistic is that measurements are consistent and that h is equal to zero.

This test statistic is equal to the objective function J from eq 23 (*5*), and thus if $\mathbf{P}_\delta^{-1}$ is a diagonal matrix, h can be rewritten as

$$h = J = \hat{\boldsymbol{\delta}}^{\mathbf{T}}\mathbf{P}_{\delta}^{-1}\hat{\boldsymbol{\delta}} = \sum_{j=1}^{n}\frac{\hat{\delta}_j^2}{\sigma_{\delta_{j,j}}^2} \quad (31)$$

with n being the number of elements in the vector $\boldsymbol{a}_{\mathrm{m}}$, i.e., the number of measured fluxes. As the error δ on the measurement is assumed to be normally distributed, the terms of the sum in the equation above are normally distributed with variance equal to one and thus the test statistic h follows, by definition, a χ^2 distribution.

It can be proven that the number of degrees of freedom of that χ^2 distribution is equal to n minus the rank of $\mathbf{R}$.

2.3.8. ***Vector Comparison Test.*** If the quality of the measurements test rejects the H_0 hypothesis that the errors are equal to zero, one can detect which measurement(s) is erroneous. A simple approach would be to try to remove measurements one by one and check if the statistical test passes (the serial elimination method of ref 5), but this is not statistically good practice as there are some issues with independency of the tests (*6*). A better approach is to use the vector comparison test (*6*).

Equation 20 can be written more explicitly as

$$\boldsymbol{R_1}a_{\mathrm{m}_1} + \boldsymbol{R_2}a_{\mathrm{m}_2} + \cdots + \boldsymbol{R_i}a_{\mathrm{m}_i} + \cdots + \boldsymbol{R_n}a_{\mathrm{m}_n} = \boldsymbol{\epsilon} \quad (32)$$

with $\boldsymbol{R_i}$ representing the different columns of $\mathbf{R}$ and a_{m_i} the different measured fluxes. If one of the measurements is really wrong by an amount τ, the expected value (E) of $\boldsymbol{\epsilon}$ is

$$\begin{aligned}\mathrm{E}[\boldsymbol{\epsilon}] &= \mathrm{E}[\boldsymbol{R_1}a_{\mathrm{m}_1} + \boldsymbol{R_2}a_{\mathrm{m}_2} + \cdots + \boldsymbol{R_i}(a_{\mathrm{m}_i} + \tau) + \cdots + \boldsymbol{R_n}a_{\mathrm{m}_n}] \\ &= \mathrm{E}[\boldsymbol{R_1}a_{\mathrm{m}_1} + \boldsymbol{R_2}a_{\mathrm{m}_2} + \cdots + \boldsymbol{R_i}a_{\mathrm{m}_i} + \cdots + \boldsymbol{R_n}a_{\mathrm{m}_n}] + \\ &\qquad \boldsymbol{R_i}\tau \\ &= \boldsymbol{R_i}\tau\end{aligned}$$

The same reasoning can be applied when more measured fluxes are wrong: the residual vector $\boldsymbol{\epsilon}$ will then be a linear combination of those erroneous measurements. For each combination of subvectors $\mathbf{R_s}$ of the redundancy matrix $\mathbf{R}$, a statistical test is run with H_0 hypothesis that the residual vector $\boldsymbol{\epsilon}$ is a linear combination of $\mathbf{R_s}$. The statistic used for this test

$$h_{\Delta} = \boldsymbol{\epsilon}^{\mathbf{T}}\mathbf{P}_{\epsilon}^{-1}\boldsymbol{\epsilon} - \boldsymbol{\epsilon}^{\mathbf{T}}\mathbf{P}_{\epsilon}^{-1}\mathbf{R_s}(\mathbf{R_s}^{\mathbf{T}}\mathbf{P}_{\epsilon}^{-1}\mathbf{R_s})^{-1}\mathbf{R_s}^{\mathbf{T}}\mathbf{P}_{\epsilon}^{-1}\boldsymbol{\epsilon} \quad (33)$$

follows a χ^2 distribution with rank($\mathbf{R}$) − rank($\mathbf{R_s}$) degrees of freedom (*6*). Running the test for each possible combination of vectors of $\mathbf{R}$ is time-consuming, but it makes no sense to take combinations with more than rank($\mathbf{R}$) − 1 vectors, as the degree of freedom of the corresponding statistic is then zero or less. This is also intuitive: a vector of n elements is a linear combination of every independent set of n vectors of length n. As such, the problem of too many combinations becomes worse the more fluxes measured, a luxury problem.

2.4. Implementation. Solving of and performing the statistical tests on metabolic models was implemented in SciLab (www.scilab.org). Figure generating was done using a combination of R (www.r-project.org), xfig (www.xfig.org) and Latex (www.latex-project.org).

All fluxes were expressed in mol L^{-1} h^{-1}, and model calculations were performed with those units. In the figures, all fluxes are divided by the biomass flux and are thus expressed as fractions of that flux.

3. Metabolic Model

3.1. Biomass Composition. To investigate whether biomass composition varied for different dilution rates, the elemental composition of the biomass was determined (the elements C, H, O, N, and S were measured; three repeats were performed). A significant difference was only found for oxygen and to a lesser extent nitrogen (Figure 1). The oxygen content increases with increasing growth rate, which is the reason the total molecular mass (expressed in C-moles) is rising slightly with increasing growth rate (Figure 2).

Measurements of the protein, DNA, and RNA content were performed to see whether the change in oxygen content was reflected in the biomass composition. However, no significant differences were found. Furthermore, sensitivity of the model to changes in the biomass composition was low. Therefore a constant biomass composition was used for every dilution rate: 70% protein, 12% RNA, 3% DNA, and 15% other (expressed in g/gDW). Those values were based partially on the values found in the literature (*25*), partially on the elemental composition of the biomass shown in Figure 1, and partially on the biomass composition measurements (data not shown).

The latter fraction is subdivided into four other components: lipopolysaccharides (LPSs), lipids, peptidoglycane, and glycogen. The relative occurrence on weight basis of those four compounds was 0.520 lipids, 0.194 LPS, 0.143 peptidoglycane, and 0.143 glycogen. Those values were taken from the average biomass composition of *E. coli* at a growth rate of 1 h^{-1} (*26*).

The composition of the lipid fraction can be found for different *E. coli* strains at different growth rates (*26*). The variations seems rather small, so constant composition was taken: 75% phosphatidylethanolamine, 20% phosphatidylglycerol, and 5% cardiolipin.

For DNA each nucleotide was assumed to be equally represented in molar units (*26*). For RNA the distribution was 26.2% ATP, 32.2% GTP, 20.0% CTP, and 21.6% UTP in molar units (*26*).

The biomass used in the metabolic model had the following elemental composition: $C_1H_{1.91}O_{0.506}N_{0.252}P_{0.015}S_{0.007}$, which gives a molecular mass of 26.22. This is somewhat higher as what would be expected from Figure 2, but it was the best agreement that could be obtained between literature data of (*25*) and the elemental composition data of Figure 1.

3.2. Different Reactions. The metabolic model included glycolysis, with glucose transported by the PTS system (*27*), the pentose phosphate pathway, the Krebs cycle, and ethanol, acetate and formate formation.

It is generally assumed that the glyoxylate pathway is not active in the *E. coli* K12 family (from which the strain used in this study is a member) if glucose is provided as a carbon source (*28*, *29*, *30*). Instead PEP carboxylase was used as regenerating reaction for the Krebs cycle metabolites (*31*, *32*). The difference between both pathways is that PEP carboxylase uses one more ATP (*32*).

For each amino acid and nucleotide the anabolic reactions were included. To avoid parallel pathways (and thus parts of the model that cannot be solved), no "salvage" pathways were used. Biosynthesis of LPS, Lipid A, peptidoglycane, and the lipid bilayer were also added. As explained above, a simple biomass composition was used for every growth rate.

Shikimic acid is an intermediate metabolite in the aromatic acid pathway, but not only shikimic acid is excreted. Also dehydroquinic acid (precursor of shikimic acid), quinic acid, dehydroshikimate, protocatechuate, and gallic acid can be excreted (*33*). Those last two were not detected in the fermentation broth and thus were not included in the model.

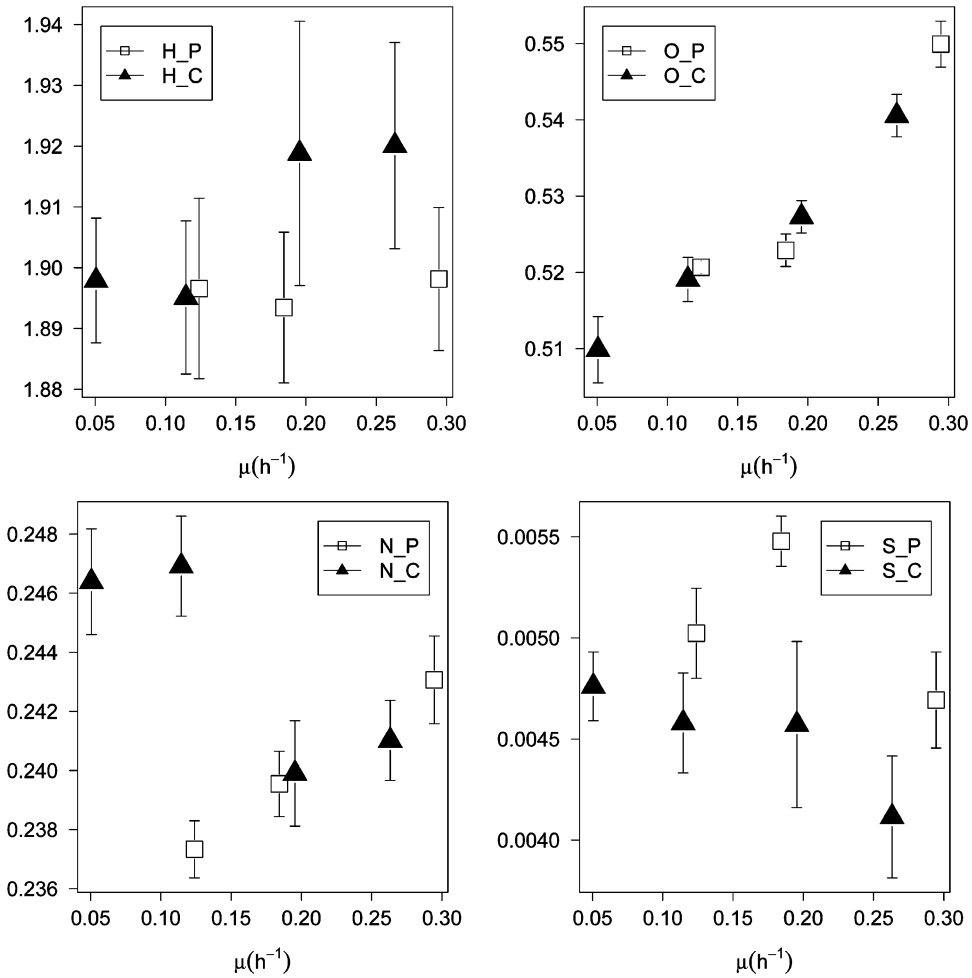

Figure 1. Elemental composition of biomass (expressed in C-moles) at different dilution rates. H (upper left), O (upper right), N (lower left) and S (lower right): (▲) C-limited cultures, (□) P-limited cultures. Error bars are for the standard deviation.

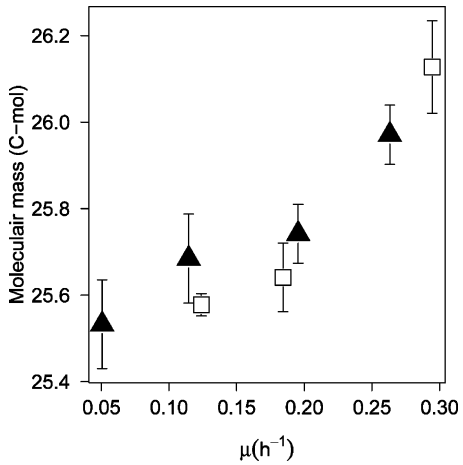

Figure 2. Molar mass of the biomass (in C-moles) as function of the dilution rate. Notation as in Figure 1.

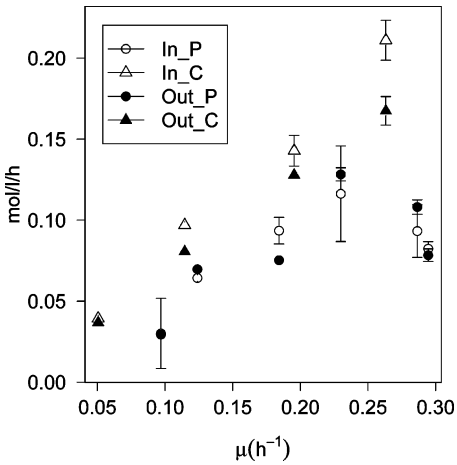

Figure 3. Moles of carbon going in the reactor versus moles of carbon leaving the reactor. Open symbols are for influxes, and closed symbols are for the carbon leaving the reactor: (△, ▲) C-limited cultures, (○, ●) P-limited cultures. Error bars represent the standard deviation.

Sources for the reactions were mainly the ecocyc database (http://www.ecocyc.org/) (*34*), the database provided by the University of California (http://systemsbiology.ucsd.edu/organisms/ecoli.html) (*35*), and the KEGG database (http://www.genome.ad.jp/kegg/) (*36*). The P/O ratio was set to 1.33 (*37*).

The thus constructed model contained 137 reactions and 151 metabolites of which 16 were exchangeable: Dhq, H_2SO_4, Ac, Cit, Shi, H_2O, O_2, PiOH, GLC, CO_2, Eth, Qa, Biom, Dhs, NH_3, and FA. All parallel pathways were removed. There were no dead-end reactions, and the elemental consistency test was passed (*20*).

There are 143 independent equations in the model and 137 + 16 unknowns. Thus, at least 10 measurements must be performed. Twelve exchange fluxes were measured: O_2, GLC,

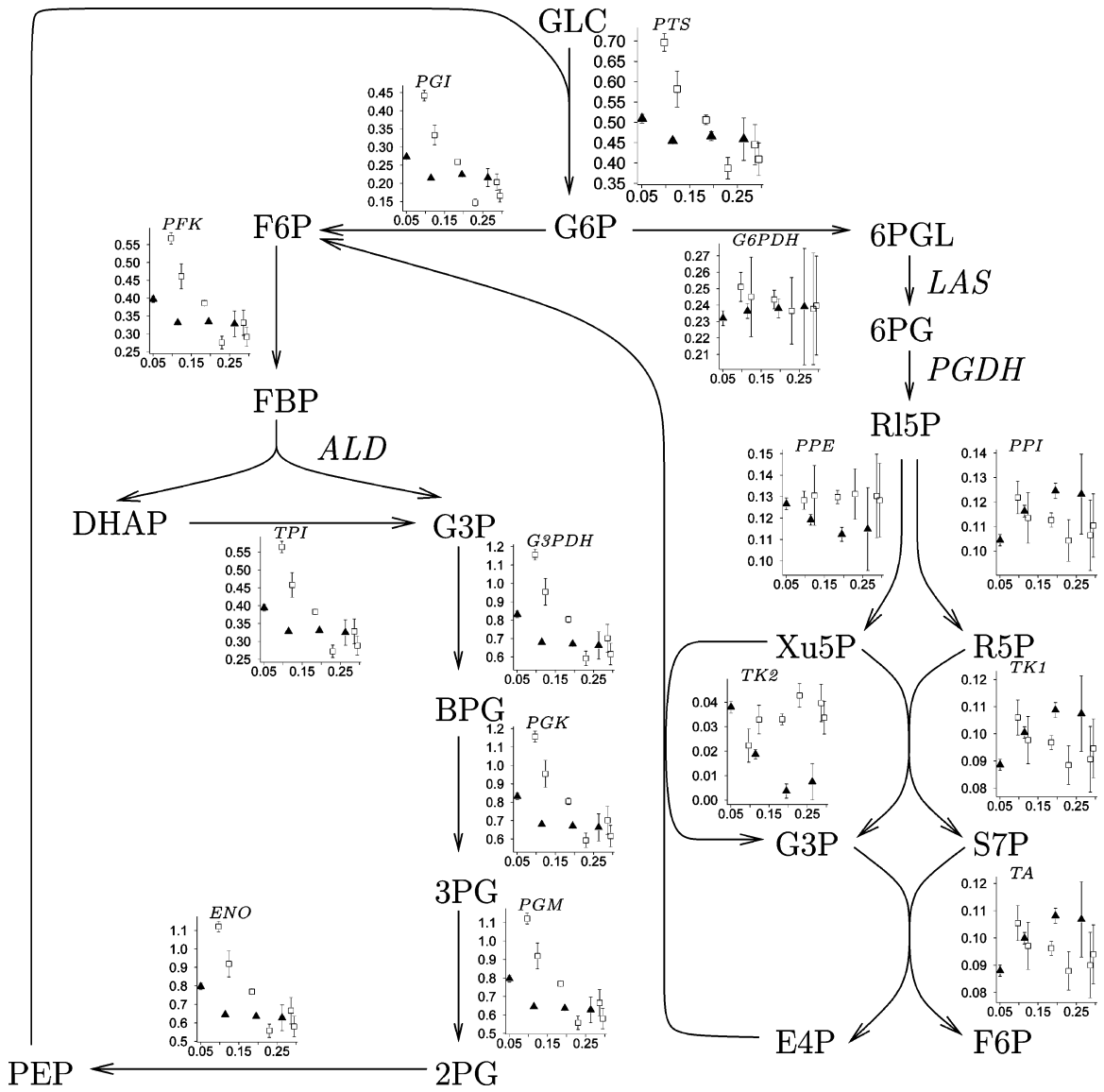

Figure 4. Relative fluxes (each flux is normalized against the biomass flux) in the glycolysis and pentose phosphate pathway for diffrerent growth rates: (□) phosphate-limited cultures, (▲) carbon-limited cultures. Metabolites are typeset in roman font, and names of reactions are in italic. The error bars represent the standard deviation.

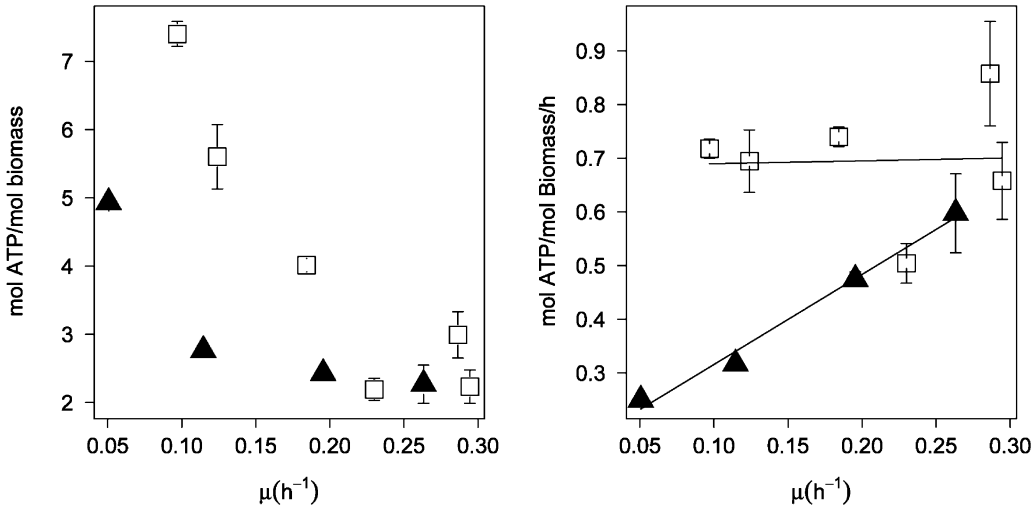

Figure 5. (Left) Moles of ATP hydrolyzed per mole of biomass formed. (Right) Moles of ATP per mole of biomass per hour that is hydrolyzed. From this data, the maintenance coefficients are calculated, assuming a constant P/O ratio of 1.l33. Notation as in Figure 4.

CO_2, Cit, FA, Shi, Dhs, Dhq, Qa, Ac, Eth, and Biom. The consistency of the measurements could thus be assessed, and if the consistency test failed, the vector comparison test could be used to search for the dissonant measurement.

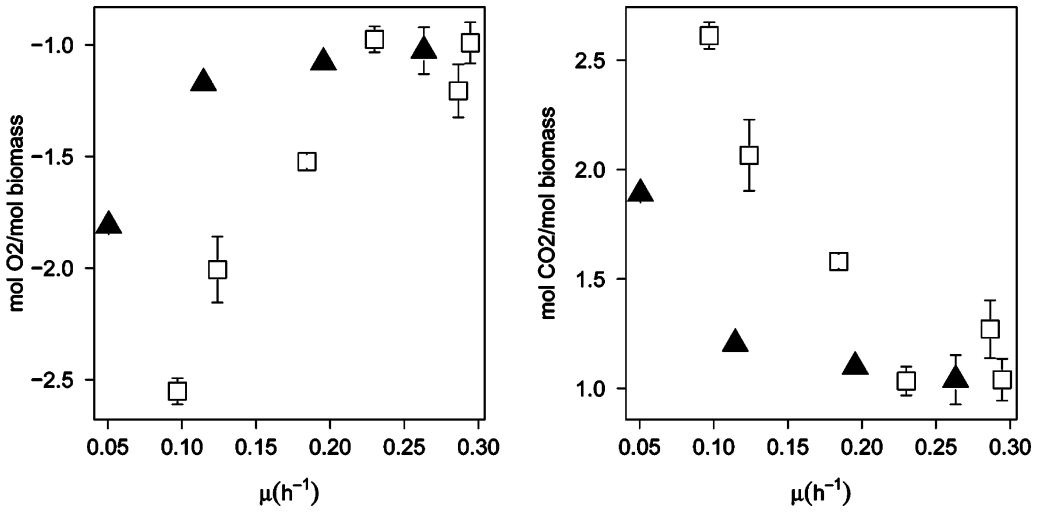

Figure 6. (Left) Moles of oxygen consumed and (right) moles of carbon dioxide produced per mole of biomass formed. Notation as in Figure 4.

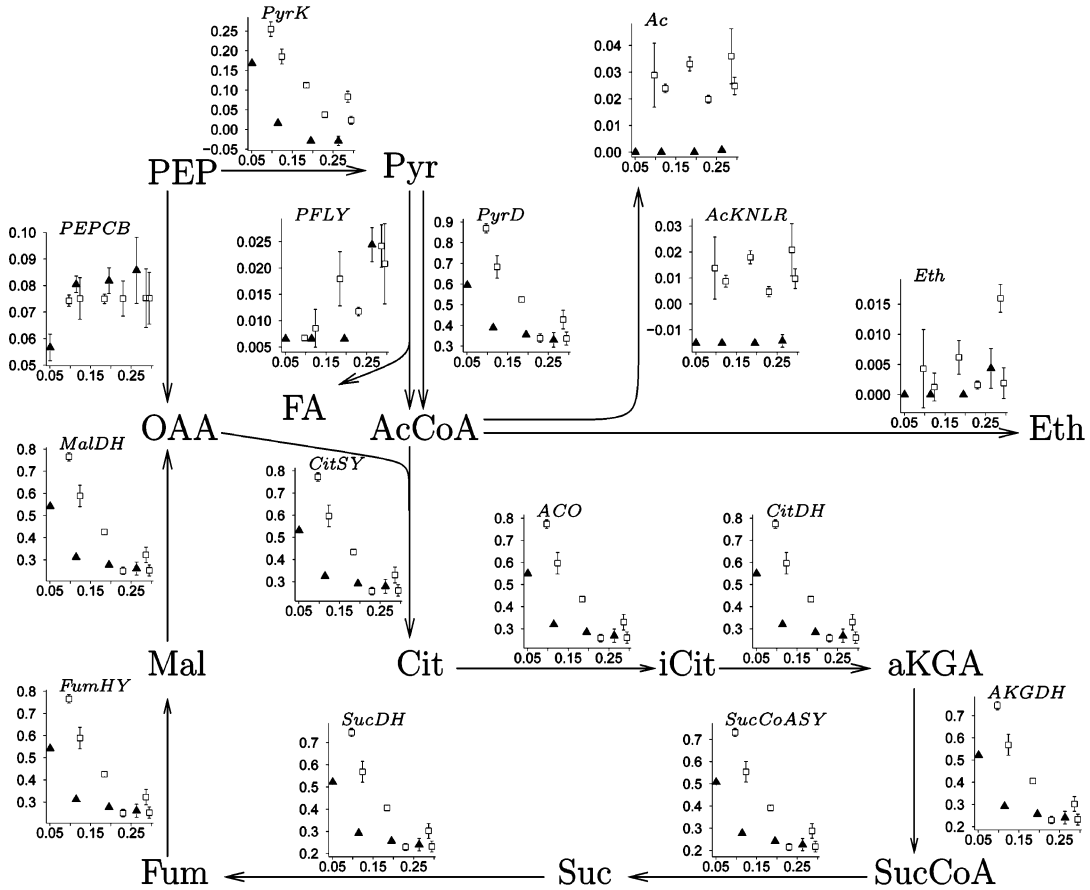

Figure 7. Fluxes in the Krebs cycle and the fermentative pathways. Notation as in Figure 4.

4. Results and Discussion

4.1. Carbon Balance. Only one statistically significant difference could be detected between the carbon in- and outflux (Figure 3): the carbon balance of the carbon-limited culture with a dilution rate of 0.12 h^{-1} is clearly not closed. It is also for this experiment that the model did not fit and no plausible wrong measurement could be detected with the vector comparison test. However, this experiment was kept in the data set since the absolute error is small even though the coincidental small variance on the measurements makes this error significant. Furthermore, the calculated fluxes were in accordance with those at other dilution rates.

4.2. Flux Analysis. Ten chemostat experiments were performed in which 12 exchange rates were measured. As explained above two measurements were redundant. The statistical test to evaluate whether the measurements agree could be run, and even when the measurements did not agree, the vector comparison test could point to the wrong measurement.

In four out of 10 experiments, the measurements were not consistent. Running the vector comparison test revealed in one case that the consumed amount of oxygen was too high and in a second case the amount of measured carbon dioxide. Removing those measurements in the respective data sets made the remaining data consistent.

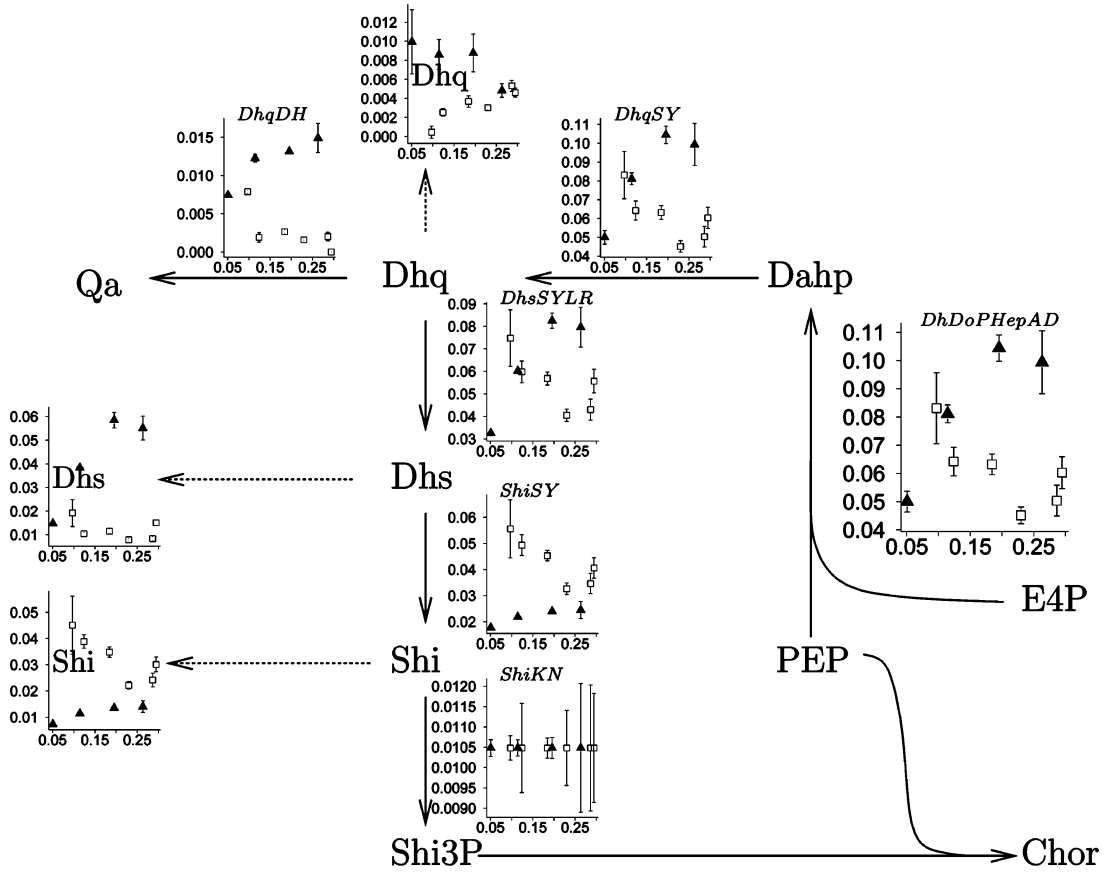

Figure 8. Fluxes in the shikimate pathway. Notation as in Figure 4.

In a third experiment, the ethanol measurement was detected as inconsistent, but removing this one made the model predict a high ethanol flux, which was never detected in reality for any of the dilution rates. However, in this dataset the oxygen measurement was identified as the second possible cause for the nonfit of the model. Removing the oxygen measurement and keeping the ethanol measurement yielded a consistent dataset and realistic model predictions.

The fourth case (the carbon-limited case with growth rate around 0.12 h^{-1}, for which the carbon balance did not close) had only the ethanol measurement as potentially erroneous. Removing that measurement made the model predict an ethanol flux that was significantly not zero. Therefore, in the simulations shown, the ethanol measurement was not removed. The data from this experiment were not discarded as they appear to agree with the cultures run at other dilution rates.

Figure 4 shows the relative fluxes (in Figures 4, 7, and 8 each molar flux is divided by the biomass flux) for the glycolysis and the pentose phosphate pathway. It can be seen that for the carbon-limited cultures, each 1 mol of glucose gives 2 mol of biomass (the graph for the PTS reaction shows that the relative flux—relative, as each molar flux is divided by the biomass flux—of glucose going into the cell is 0.5, thus to produce 1 mol of biomass, 0.5 mol of glucose is consumed), or put another way, three of the six carbons of glucose are used for biomass construction. The same graph also shows that under P-limited conditions, at low dilution rates, the biomass yield is lower, but strangely, at a dilution rate of 0.25 h^{-1} it tends to become better than under carbon limitation (less glucose is consumed per mole of biomass produced than under carbon limitation). Looking at the ATP hydrolysis (Figure 5) and the respiration (Figure 6) there is no difference between the carbon-limited and phosphate-limited cultures at higher dilution rates. Thus, the better biomass yield of P-limited cultures at high dilution rates is not due to less CO_2 production or less ATP hydrolysis. The reason for it should be found in the aromatic pathway.

In the first reaction of the aromatic pathway, the synthesis of Dahp (Figure 8), one can see that for high dilution rates, the flux through it is higher for carbon-limited cultures than for phosphate-limited ones. Considering that the fluxes through the chorismate synthesis reaction are all equal (because a fixed biomass composition was used), the increased flux through the Dahp formation reaction necessarily means an increased excretion of products upstream of chorismate. Unfortunately, those byproducts were found not to be shikimic acid, the target molecule, but dehydroquinate and dehydroshikimate. Thus, carbon-limited cultures have a larger flux through the shikimate pathway but do not produce more shikimate because that extra flux is wasted on other byproducts. If this byproduct excretion could be eliminated, carbon-limited cultures would clearly be better for shikimate production.

The larger flux through the shikimate route for carbon-limited cultures gives no significant rise in the G6P levels that enter the pentose phosphate pathway. The differentiation is in the amount of carbon that goes from the PPP to the glycolysis pathway via F6P and G3P. A split occurs at Rl5P. More E4P is needed to sustain the production of Dahp under carbon limitation, and therefore the reaction catalyzed by TK2 is lower or even zero at higher dilution rates for carbon-limited cultures. This makes more Xu5P available, which in turn allows the TK1 flux to increase. This allows the TA flux to increase, leading to a higher E4P production. As under carbon limitation the flux through TK2 is lower, the one through PPE can also be lower, whereas the one through PPI is higher. This is in line with the studies where an increased flux to the shikimate pathway was achieved by overexpression of Tk1. An amplification of this

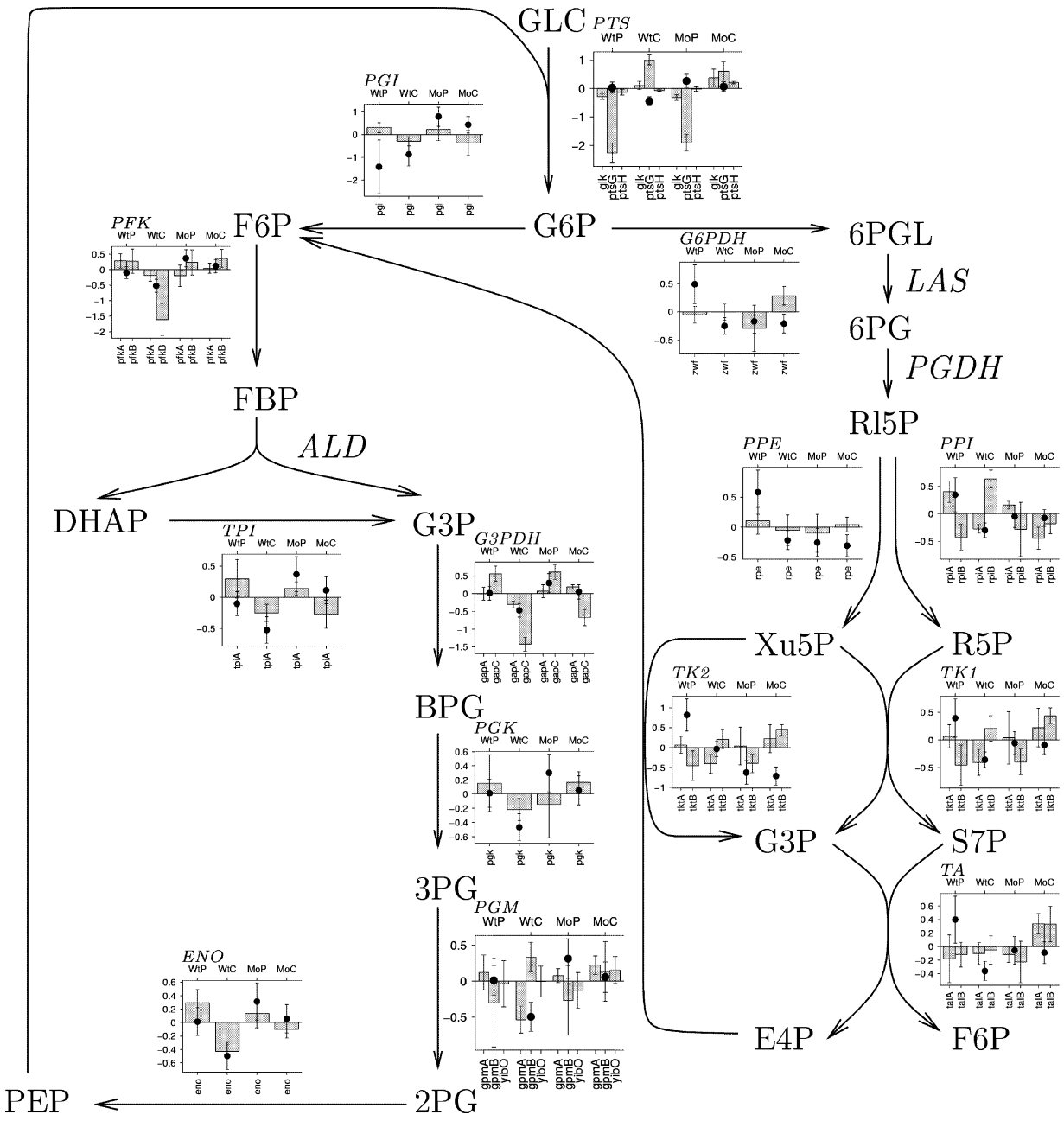

Figure 9. Fluxes compared with the RNA expression data for the central carbon metabolism. The error bars represent the standard deviation. The dots represent the fluxes, and the bars represent the gene expression levels. The gene names can be found in the lower part of each graph, and the upper part labels the different experiments: WtP for wild-type strain, P-limited culture; WtC for wild-type strain, C-limited culture; MoP for modified strain, P-limited culture; and MoC for modified strain, C-limited culture. Reaction names are typeset in italic font, and metabolite names are in roman font.

enzyme in combination with an overexpressed DAHP-synthase has shown to double the flow into the pathway in comparison to the case where only DAHP-synthase was overexpressed (*38*).

The fluxes through glycolysis (Figure 4) follow the same pattern as those from ATP hydrolysis (Figure 5) and respiration (Figure 6). This pathway is mainly used to fuel the citric acid cycle (Figure 7) for generating ATP and biomass precursors. Interesting to note is the negative flux in carbon-limited cultures going through AcKNLR (Figure 7). Actually, no acetate was supplied to the medium. The acetate consumed is provided by the cysteine, ornithine, and lipid A synthesis reactions. Acetate consumption in the presence of glucose, although uncommon, was reported in the literature (*39*).

The flux through PEP carboxylase (Figure 7) is completely dependent on the biomass formation reaction (as is the one through the chorismate synthesis reaction, Figure 8). This can be seen nicely as the constant flux for every phosphate-limited culture. For the carbon-limited cultures, this is not true, as there was some citric acid consumed in the experiment with the lowest growth rate, whereas it was produced in the other ones. The amounts of citric acid consumed or produced were however very low: there is no significant difference between the carbon-limited and phosphate-limited cultures in the flux through PEP carboxylase.

The calculated ATP hydrolysis flux allows to investigate the maintenance requirement for different growth rates and limiting conditions (Figure 5). Phosphate-limited cultures have a higher non-growth-associated maintenance compared to that of carbon-limited ones as under carbon abundance there is no need for an effective carbon utilization. The non-growth-associated main-

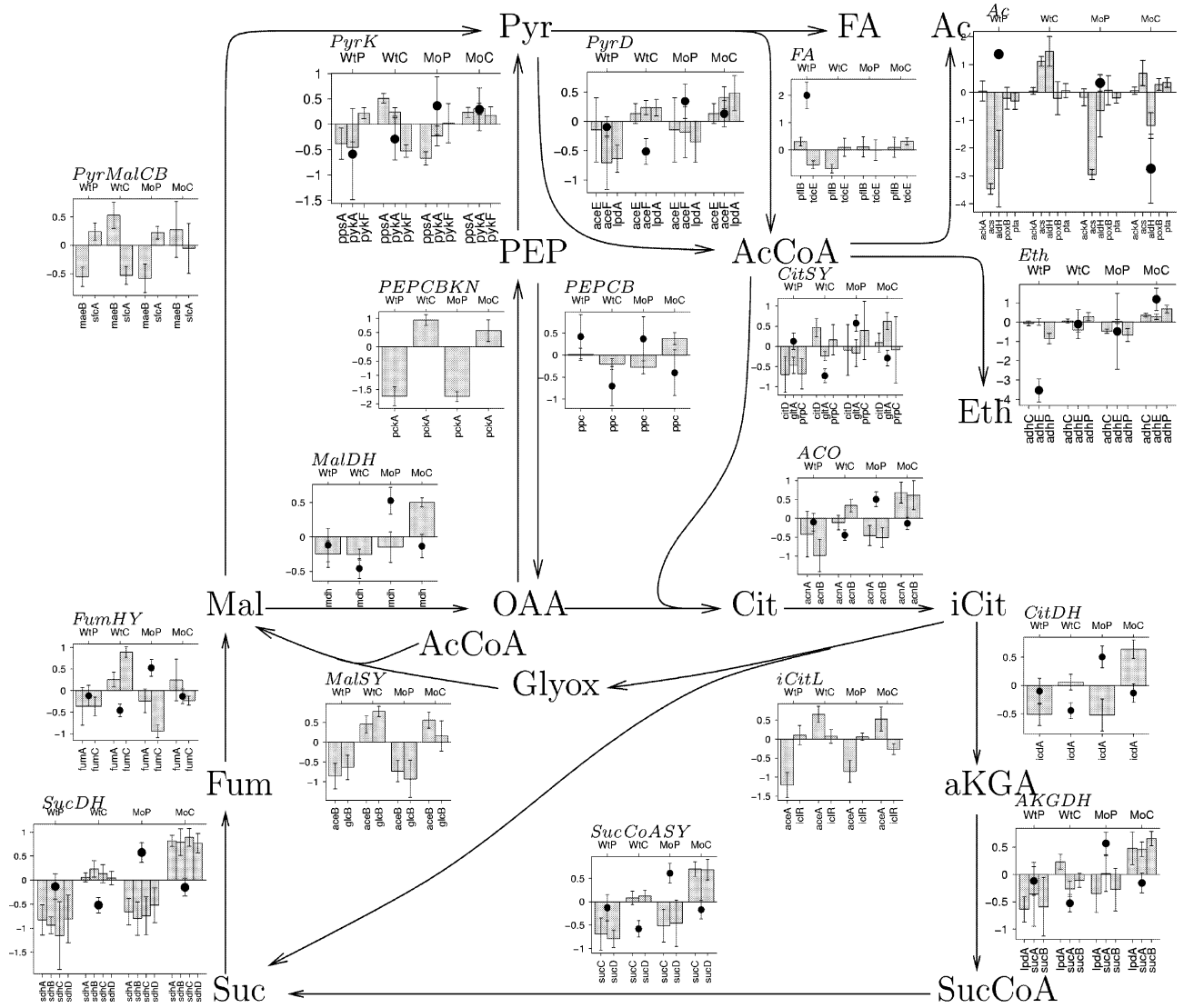

Figure 10. Fluxes compared with the RNA expression data for the Krebs cycle. Notation as in Figure 9.

tenance for phosphate-limited cultures is 0.68 mol ATP per mole of biomass per hour. For carbon-limited cultures it is 0.15 mol/molBM/h. This conforms with the values reported in the literature: 0.20 mol/molBM/h (*39*), 0.073 mol/molBM/h (*40*), and 0.12 mol/molBM/h (*41*).

The growth-associated maintenance for phosphate-limited cultures is very low: 0.053 mol per mole of biomass. For carbon-limitation this is 1.68 mol/molBM/h. In literature both low values, 0.34 mol/molBM (*39*), and high values, 2.6 mol/molBM (*41*) can be found.

Apparently under carbon-abundant conditions, the cells, even at the lowest growth rate, produce as much ATP as possible and this production does not significantly increase (cannot increase?) with higher growth rates. Carbon-limited cultures want to optimize the utilization of the available glucose and try to minimize the maintenance cost, resulting in an high growth dependent maintenance. However, the total maintenance is always lower than in carbon-abundant cultures (in the right part of Figure 5, the curve of the P-limited cultures, squares, lies above the curve of the C-limited ones, triangles). At a growth rate of 0.3 h^{-1} both curves meet each other. This is also the point at which the cells start to wash out. This suggests that the maintenance level of the phosphate-limited cultures is really the maximum that the cell can sustain.

4.3. RNA Expression Compared with Metabolic Fluxes. To investigate whether there exists a correlation between RNA expression levels and metabolic fluxes, four different kinds of cultures were performed in which flux data and RNA expression levels were determined. Two different experiments were done with the wild-type strain (no genetic modifications for increased flux through the aromatic pathway), one carbon-limited and one phosphate-limited, and two experiments were done with the modified strain, also one carbon-limited and one phosphate-limited. Each experiment was performed twice. The continuous cultures were all run at a dilution rate of 0.2 h^{-1}.

In general RNA expression levels are compared by taking the $\log_2$ of the fraction of expression levels under modified conditions against expression levels in a reference state. The reference state in this case would be the wild-type fermentation under carbon- or phosphate-limited conditions. As the RNA expression levels have to be compared with the fluxes, the same mathematical treatment should be applied to the flux values (relative molar fluxes were used; each flux expressed in mol L^{-1} h^{-1} was divided by the biomass flux). However, for the wild type, some fluxes in the shikimate-producing pathway are zero (more precisely, there is no quinic acid production). For the modified strain, some other fluxes are zero (no formic acid formation). Thus, for those fluxes, no good reference state exists.

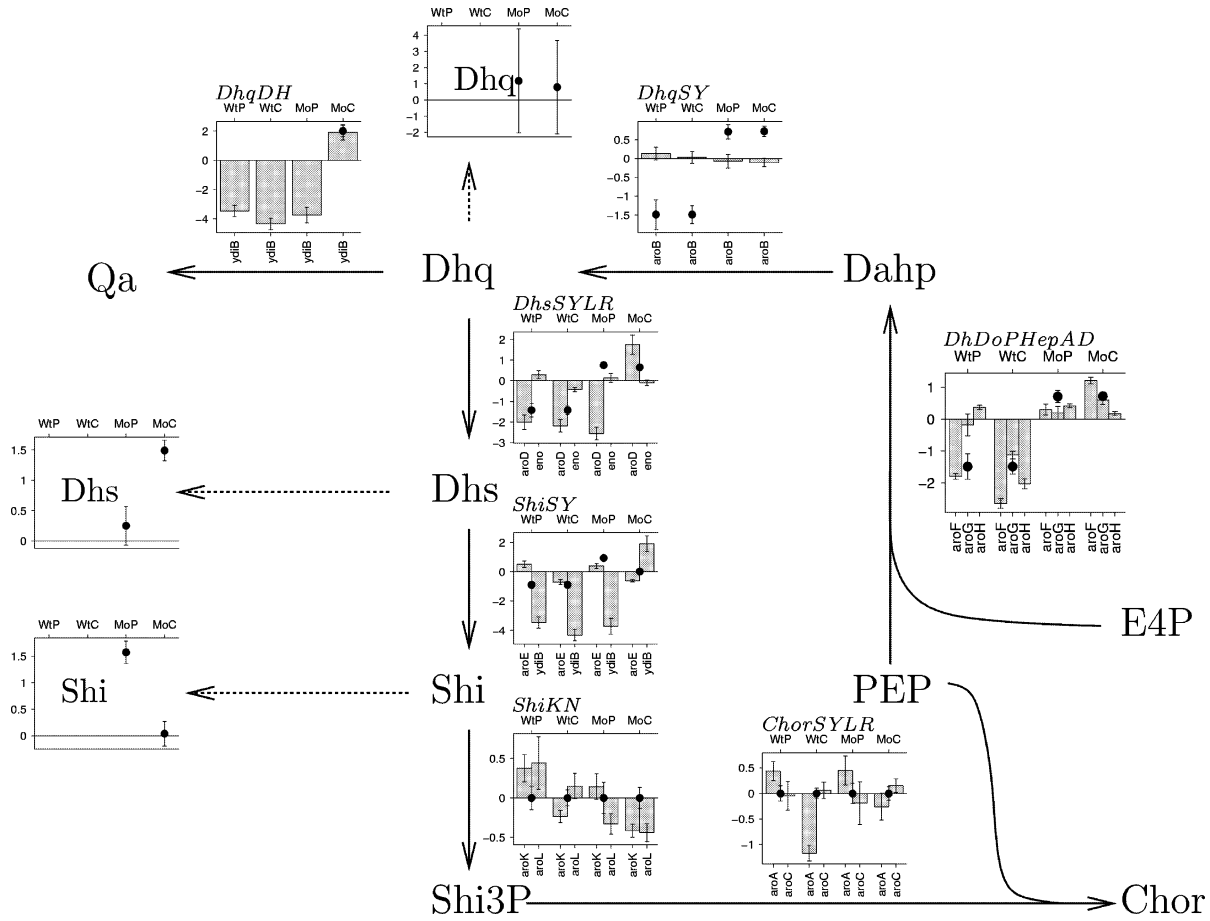

Figure 11. Fluxes compared with the RNA expression data in the aromatic amino acid pathway. Notation as in Figure 9.

Therefore, the mean of the four values was used as reference state, each value was divided by it, and the $\log_2$ was taken. This was done for both the RNA expression levels and the fluxes that were not zero (Figures 9−11).

The flux through a reaction is not always controlled by the enzyme level, as metabolic control analysis shows, but can also be limited by the availability of reactants. It is not expected that there is a direct correlation between RNA expression levels and enzyme levels, even less between RNA expression levels and metabolic fluxes. Only when RNA expression is absent can one say for sure that the corresponding flux is zero, and the absence of expression of a gene can only be disproven, not proven (*12*). However, in some cases there can be qualitative correlations between RNA expression levels and metabolic fluxes (*42*). For some (key) reactions an increase in flux correlates with an increase or decrease in RNA expression levels (Figures 9−11). Some correlations were detected in the glycolysis and the aromatic pathway, whereas in the PPP and the citric acid cycle the reactions seem to be driven by supply and demand and not by RNA expression of their genes.

In the following paragraphs examples are given of these three cases: positive correlation, negative correlation and no correlation.

4.3.1. ***Positive Correlation.*** The modified strain has the feedback inhibition acting on *aroF* and *aroG* from tyrosine and phenylalanine, respectively, trypthophane and phenylalanine removed. This can be seen both in the fluxes and in the RNA expression levels (Figure 11).

The expression of *ydiB* seems not to be correlated with the ShiSY reaction (Figure 11), but it is correlated with the production of quinate. Carbon limitation in the modified strains gives a high *ydiB* expression combined with a high flux through DhqDH.

4.3.2. ***Negative Correlation.*** Whereas in C-limited cultures *ptsG* is more expressed than in carbon-abundant ones, the specific uptake of glucose is lower (Figure 9). An explanation might be that under such conditions the cell tries to optimize the uptake of carbon by expressing more *ptsG* (*43*). The same is valid for the gene *acs*, for acetate assimilation (Figure 10): this gene codes for acetyl coenzyme A synthethase and is even expressed when no acetate is available (*44*). This seems to fit the case here: under carbon-limiting conditions, the expression of this gene is stronger; more need for carbon?

Carbon-abundant cultures have a low expression of genes coding for pyruvate dehydrogenase (*PyrD* in Figure 10) and a high flux through that reaction, whereas carbon-limited cultures show the opposite behavior. This confirms the high non-growth- and low growth-associated maintenance found for the phosphate-limited cultures: the flux through the citric acid cycle (mainly used for generating ATP via respiration) is saturated. This can also be observed in the Krebs cycle itself: the flux through it is higher in P- than in C-limitation while the genes are downregulated. Furthermore, the flux in the modified strain is generally higher than the one in the wild-type strain.

4.3.3. ***No Correlation.*** In general the fluxes in the PPP pathway do not follow the RNA expression patterns (Figure 9). Fluxes seem to be driven by the need to build blocks for the amino acid synthesis reactions, in case of the aromatic pathway.

PGI: G6P ⟷ F6P
PFK: ATP + F6P ⟶ ADP + FBP
ALD: FBP ⟷ G3P + DHAP
TPI: DHAP ⟷ G3P
G3PDH: PiOH + NAD + G3P ⟷ NADH + H + BPG
PGK: ADP + BPG ⟷ ATP + 3PG
PGM: 3PG ⟷ 2PG
ENO: 2PG ⟷ H2O + PEP
PyrK: ADP + PEP ⟶ ATP + Pyr
PyrD: NAD + Pyr + CoA ⟶ NADH + H + AcCoA + CO2
CitSY: H2O + AcCoA + OAA ⟶ CoA + Cit
ACO: Cit ⟷ iCit
CitDH: NAD + iCit ⟷ NADH + H + CO2 + aKGA
AKGDH: NAD + CoA + aKGA ⟶ NADH + H + CO2 + SucCoA
SucCoASY: ADP + PiOH + SucCoA ⟷ ATP + CoA + Suc
SucDH: FAD + Suc ⟶ FADH2 + Fum
FumHY: H2O + Fum ⟷ Mal
MalDH: NAD + Mal ⟷ NADH + H + OAA
PEPCB: H2O + PEP + CO2 ⟶ PiOH + OAA
PFLY: Pyr + CoA ⟶ AcCoA + FA
EthDHLR: 2 NADH + 2 H + AcCoA ⟷ 2 NAD + CoA + Eth
AcKNLR: ADP + PiOH + AcCoA ⟷ ATP + CoA + Ac
Resp: 1.33 ADP + 1.33 PiOH + NADH + H + 0.5 O2 ⟶ 1.33 ATP + NAD + 2.33 H2O
H2CO3SY: H2O + CO2 ⟷ H2CO3
G6PDH: NADP + G6P ⟶ NADPH + H + 6PGL
LAS: H2O + 6PGL ⟶ 6PG
PGDH: NADP + 6PG ⟶ NADPH + H + CO2 + Rl5P
PPI: Rl5P ⟷ R5P
PPE: Rl5P ⟷ Xu5P
TK1: R5P + Xu5P ⟷ G3P + S7P
TA: G3P + S7P ⟷ F6P + E4P
TK2: Xu5P + E4P ⟷ F6P + G3P
PTS: GLC + PEP ⟶ G6P + Pyr
PPiOHHY: PPiOH + H2O ⟶ 2 PiOH
GluDH: NADPH + H + aKGA + NH3 ⟷ NADP + H2O + Glu
GluLI: ATP + NH3 + Glu ⟶ ADP + PiOH + Gln
AspSY: ATP + H2O + Asp + Gln ⟶ AMP + PPiOH + Asn + Glu
AspTA: OAA + Glu ⟷ aKGA + Asp
AlaTA: Pyr + Glu ⟷ aKGA + Ala
ValAT: aKIV + Glu ⟷ aKGA + Val
LeuSYLR: NAD + H2O + AcCoA + aKIV + Glu ⟶ NADH + H + CoA + CO2 + aKGA + Leu
aKIVSYLR: NADPH + H + 2 Pyr ⟶ NADP + H2O + CO2 + aKIV
IleSYLR: NADPH + H + Pyr + Glu + Thr ⟶ NADP + H2O + CO2 + aKGA + NH3 + Ile
ProSYLR: ATP + 2 NADPH + 2 H + Glu ⟶ ADP + PiOH + 2 NADP + H2O + Pro
SerLR: NAD + H2O + 3PG + Glu ⟶ PiOH + NADH + H + aKGA + Ser
SerTHM: Ser + THF ⟶ H2O + Gly + MeTHF
H2SSYLR: 2 ATP + 3 NADPH + ThioredH2 + 3 H + H2SO4 ⟶ ADP + PPiOH + 3 NADP + Thiored + 3 H2O + H2S + PAP
PAPNAS: H2O + PAP ⟶ AMP + PiOH
CysSYLR: H2S + AcCoA + Ser ⟶ CoA + Cys + Ac
PrppSY: ATP + R5P ⟶ AMP + PRPP
HisSYLR: ATP + 2 NAD + 3 H2O + Gln + PRPP ⟶ 2 PPiOH + PiOH + 2 NADH + 2 H + aKGA + His + AICAR
PheSYLR: Glu + Chor ⟶ H2O + CO2 + aKGA + Phe
TyrSYLR: NAD + Glu + Chor ⟶ NADH + H + CO2 + aKGA + Tyr
TrpSYLR: Gln + Ser + Chor + PRPP ⟶ PPiOH + 2 H2O + G3P + Pyr + CO2 + Glu + Trp
DhDoPHepAD: H2O + PEP + E4P ⟶ PiOH + Dahp
DhqSY: Dahp ⟶ PiOH + Dhq
DhsSYLR: Dhq ⟷ H2O + Dhs
ShiSY: NADPH + H + Dhs ⟷ NADP + Shi
ShiKN: ATP + Shi ⟶ ADP + Shi3P
DhqDH: NADPH + H + Dhq ⟶ NADP + Qa
ChorSYLR: PEP + Shi3P ⟶ 2 PiOH + Chor
ThrSYLR: ATP + H2O + HSer ⟶ ADP + PiOH + Thr
MDAPSYLR: NADPH + H + Pyr + SucCoA + Glu + AspSA ⟶ NADP + CoA + aKGA + Suc + MDAP
LysSY: MDAP ⟶ CO2 + Lys
MetSYLR: H2O + SucCoA + Cys + MTHF + HSer ⟶ Pyr + CoA + Suc + NH3 + Met + THF
AspSASY: ATP + NADPH + H + Asp ⟶ ADP + PiOH + NADP + AspSA
HSerDH: NADPH + H + AspSA ⟷ NADP + HSer
CarPSY: 2 ATP + H2O + H2CO3 + Gln ⟶ 2 ADP + PiOH + Glu + CarP
OrnSYLR: ATP + NADPH + H + H2O + AcCoA + 2 Glu ⟶ ADP + PiOH + NADP + CoA + aKGA + Orn + Ac
ArgSYLR: ATP + Asp + Orn + CarP ⟶ AMP + PPiOH + PiOH + Fum + Arg
ThioredRD: NADPH + Thiored + H ⟷ NADP + ThioredH2
H2O2ox: 2 H2O2 ⟶ 2 H2O + O2
FAD2NAD: NAD + FADH2 ⟷ NADH + FAD + H
AICARSYLR: 6 ATP + 3 H2O + CO2 + Asp + 2 Gln + Gly + FA + PRPP ⟶ 6 ADP + PPiOH + 6 PiOH + Fum + 2 Glu + AICAR
IMPSYLR: FTHF + AICAR ⟶ H2O + THF + IMP
AMPSYLR: Asp + GTP + IMP ⟶ AMP + PiOH + Fum + GDP
AdKN: ATP + AMP ⟷ 2 ADP
ADPRD: ADP + ThioredH2 ⟶ Thiored + H2O + dADP
dADPKN: ATP + dADP ⟶ ADP + dATP
IMPDH: NAD + H2O + IMP ⟶ NADH + H + XMP
GMPSY: ATP + H2O + Gln + XMP ⟶ AMP + PPiOH + Glu + GMP
GuKN: ATP + GMP ⟶ ADP + GDP
GDPKN: ATP + GDP ⟶ ADP + GTP
GDPRD: ThioredH2 + GDP ⟶ Thiored + H2O + dGDP
dGDPKN: ATP + dGDP ⟶ ADP + dGTP
UMPSYLR: O2 + Asp + PRPP + CarP ⟶ PPiOH + PiOH + H2O + CO2 + UMP + H2O2
UrKN: ATP + UMP ⟶ ADP + UDP
UDPKN: ATP + UDP ⟶ ADP + UTP
CTPSY: ATP + H2O + Gln + UTP ⟶ ADP + PiOH + Glu + CTP
CDPKN: ATP + CDP ⟷ ADP + CTP
CMPKN: ATP + CMP ⟶ ADP + CDP
CDPRD: ThioredH2 + CDP ⟶ Thiored + H2O + dCDP
dCDPKN: ATP + dCDP ⟶ ADP + dCTP
UDPRD: ThioredH2 + UDP ⟶ Thiored + H2O + dUDP
dUDPKN: ATP + dUDP ⟶ ADP + dUTP
dUTPPPAS: H2O + dUTP ⟶ PPiOH + dUMP
dTMPSY: MeTHF + dUMP ⟶ DHF + dTMP
dTMPKN: ATP + dTMP ⟶ ADP + dTDP
dTDPKN: ATP + dTDP ⟶ ADP + dTTP
DHFRD: NADPH + H + DHF ⟶ NADP + THF
FTHFSYLR: NADP + H2O + MeTHF ⟶ NADPH + H + FTHF
GlyCA: NAD + Gly + THF ⟷ NADH + H + CO2 + NH3 + MeTHF
MeTHFRD: NADH + H + MeTHF ⟶ NAD + MTHF
AcCoACB: ATP + H2O + AcCoA + CO2 ⟷ ADP + PiOH + MalCoA
MalCoATA: MalCoA + ACP ⟷ CoA + MalACP
AcACPSY: MalACP ⟶ CO2 + AcACP
C120SY: 10 NADPH + 10 H + AcACP + 5 MalACP ⟶ 10 NADP + 5 H2O + 5 CO2 + C120ACP + 5 ACP
C140SY: 12 NADPH + 12 H + AcACP + 6 MalACP ⟶ 12 NADP + 6 H2O + 6 CO2 + C140ACP + 6 ACP
C160SY: 14 NADPH + 14 H + AcACP + 7 MalACP ⟶ 14 NADP + 7 H2O + 7 CO2 + C160ACP + 7 ACP
C181SY: 15 NADPH + 15 H + AcACP + 8 MalACP ⟶ 15 NADP + 8 H2O + 8 CO2 + C181ACP + 8 ACP
AcylTF: C160ACP + C181ACP + Go3P ⟶ 2 ACP + PA
Go3PDH: NADPH + H + DHAP ⟷ NADP + Go3P
DGoKN: ATP + DGo ⟶ ADP + PA
CDPDGoSY: CTP + PA ⟷ PPiOH + CDPDGo
PSerSY: Ser + CDPDGo ⟶ CMP + PSer
PSerDC: PSer ⟶ CO2 + PEthAn
GlnF6PTA: F6P + Gln ⟶ Glu + GA6P
GlcAnMU: GA6P ⟷ GA1P
NAGUrTF: AcCoA + UTP + GA1P ⟶ PPiOH + CoA + UDPNAG
LipaSYLR: ATP + 2 CMPKDO + 2 UDPNAG + C120ACP + 5 C140ACP ⟶ ADP + 2 CMP + UMP + UDP + 6 ACP + Lipa + 2 Ac
A5PIR: Rl5P ⟷ Ar5P
PGLCMT: G6P ⟷ G1P
CMPKDOSYLR: 2 H2O + PEP + Ar5P + CTP ⟶ PPiOH + 2 PiOH + CMPKDO
ADPHEPSY: ATP + S7P ⟶ PPiOH + ADPHEP
UDPGlcSY: G1P + UTP ⟶ PPiOH + UDPGlc
EthANPT: CMP + PEthAn ⟷ CDPEthAn + DGo
LpsSYLR: 3 ADPHEP + 3 CMPKDO + 2 UDPGlc + Lipa + 2 CDPEthAn ⟶ 3 ADP + 3 CMP + 2 CDP + 2 UDP + Lps
PGSYLR: Go3P + CDPDGo ⟶ PiOH + CMP + PG
CLSY: PG + CDPDGo ⟶ CMP + CL
PeptidoSYLR: 5 ATP + NADPH + H + PEP + 3 Ala + MDAP + 2 UDPNAG ⟶ 5 ADP + 7 PiOH + NADP + UMP + UDP + Peptido
GlcgSY: ATP + G1P ⟶ ADP + PPiOH + Glcg
ATPHY: ATP + H2O ⟶ ADP + PiOH
DNASYLR: 2 H2O + 0.246 dATP + 0.254 dGTP + 0.254 dCTP + 0.246 dTTP ⟶ 2 PiOH + DNA
RNASYLR: 0.262 ATP + 2 H2O + 0.322 GTP + 0.2 CTP + 0.216 UTP ⟶ 2 PiOH + RNA
ProtSYLR: 2 ATP + 4 H2O + 0.0961 Ala + 0.05506 Arg + 0.04505 Asn + 0.04505 Asp + 0.01702 Cys + 0.04905 Gln + 0.04905 Glu + 0.1151 Gly + 0.01802 His + 0.05405 Ile + 0.08408 Leu + 0.06406 Lys + 0.02903 Met + 0.03504 Phe + 0.04104 Pro + 0.04004 Ser + 0.04705 Thr + 0.01101 Trp + 0.02603 Tyr + 0.07908 Val + 2 GTP ⟶ 2 ADP + 4 PiOH + 2 GDP + Prot
LipidSYLR: 0.0266 CL + 0.202 PG + 0.7714 PEthAn ⟶ Lipid
BiomSYLR: 0.003472 Glcg + 0.0002027 Lps + 0.0006801 Peptido + 0.002408 DNA + 0.009191 RNA + 0.1454 Prot + 0.002774 Lipid ⟶ Biom

Figure 12. Appendix A: List of reactions included in the model.

2PG	$C_3H_7O_7P$	2-phophoglycerate
3PG	$C_3H_7O_7P$	3-phophoglycerate
6PG	$C_6H_{13}O_{10}P$	6-phosphogluconate
6PGL	$C_6H_{11}O_9P$	6-phosphogluconolacton
Ac	$C_2H_4O_2$	Acetate
AcACP	C_2H_3OPept	Acetyl ACP
AcCoA	$C_{23}H_{34}O_{17}N_7P_3S$	Acetyl CoA
ACP	HPept	Acyl carier protein
ADP	$C_{10}H_{15}O_{10}N_5P_2$	Adenosine diphosphate
ADPHEP	$C_{17}H_{27}O_{16}N_5P_2$	ADP-Mannoheptose
AICAR	$C_9H_{15}O_8N_4P$	Amino imidazole carboxamide ribonucleotide
aKGA	$C_5H_6O_5$	Alpha keto glutaric acid
aKIV	$C_5H_8O_3$	Alpha-keto-isovalerate
Ala	$C_3H_7O_2N$	Alanine
AMP	$C_{10}H_{14}O_7N_5P$	Adenosine monophosphate
Ar5P	$C_5H_{11}O_8P$	Arabinose-5-phosphate
Arg	$C_6H_{14}O_2N_4$	Arginine
Asn	$C_4H_8O_3N_2$	Aspartate
Asp	$C_4H_7O_4N$	Asparagine
AspSA	$C_4H_7O_3N$	Aspartate semialdehyde
ATP	$C_{10}H_{16}O_{13}N_5P_3$	Adenosine triphosphate
Biom	$CH_{1.91}O_{0.506}N_{0.252}P_{0.0149}S_{0.0067}$	Biomass
BPG	$C_3H_8O_{10}P_2$	1-3-biphosphoglycerate
C120ACP	$C_{12}H_{23}OPept$	
C140ACP	$C_{14}H_{27}OPept$	
C160ACP	$C_{16}H_{31}OPept$	
C181ACP	$C_{18}H_{33}OPept$	
CarP	CH_4O_5NP	Carbamoyl phosphate
CDP	$C_9H_{15}O_{11}N_3P_2$	Citidine diphosphate
CDPDGo	$C_{46}H_{83}O_{15}N_3P_2$	CDP-diacylglycerol
CDPEthAn	$C_{11}H_{20}O_{11}N_4P_2$	CDP-ethanolamine
Chor	$C_{10}H_{10}O_6$	Chorismate
Cit	$C_6H_8O_7$	cisaconitate
CL	$C_{77}H_{144}O_{16}P_2$	Cardiolipin
CMP	$C_9H_{14}O_8N_3P$	Citidine monophosphate
CMPKDO	$C_{17}H_{26}O_{15}N_3P$	CMP-2-keto-3-deoxyoctanoate
CO2	CO_2	Carbondioxide
CoA	$C_{21}H_{32}O_{16}N_7P_3S$	Coenzyme A
CTP	$C_9H_{16}O_{14}N_3P_3$	Citidine triphosphate
Cys	$C_3H_7O_2NS$	Cysteine
dADP	$C_{10}H_{15}O_9N_5P_2$	deoxy ADP
Dahp	$C_7H_{13}O_{10}P$	Deoxy arabino heptulosonate
dATP	$C_{10}H_{16}O_{12}N_5P_3$	deoxy ATP
dCDP	$C_9H_{15}O_{10}N_3P_2$	deoxy CDP
dCTP	$C_9H_{16}O_{13}N_3P_3$	deoxy CTP
dGDP	$C_{10}H_{15}O_{10}N_5P_2$	deoxy GDP
DGo	$C_{37}H_{70}O_5$	Diacyl glycerol
dGTP	$C_{10}H_{16}O_{13}N_5P_3$	deoxy GTP
DHAP	$C_3H_7O_6P$	Dihydroxyaceton phosphate
DHF	$C_{19}H_{21}O_6N_7$	Dihydrofolate
Dhq	$C_7H_{10}O_6$	Dehydroquinate
Dhs	$C_7H_8O_5$	Dehydroshikimate
DNA	$C_{9.75}H_{14.2}O_7N_{3.75}P$	DNA composition
dTDP	$C_{10}H_{16}O_{11}N_2P_2$	deoxy TDP
dTMP	$C_{10}H_{15}O_8N_2P$	deoxy TMP
dTTP	$C_{10}H_{17}O_{14}N_2P_3$	deoxy TTP
dUDP	$C_9H_{14}O_{11}N_2P_2$	deoxy UDP
dUMP	$C_9H_{13}O_8N_2P$	deoxy UMP
dUTP	$C_9H_{15}O_{14}N_2P_3$	deoxy UTP
E4P	$C_4H_9O_7P$	Erythrose-4-phosphate
Eth	C_2H_6O	Ethanol
F6P	$C_6H_{13}O_9P$	Fructose-6-phosphate
FA	CH_2O_2	Formic Acid
FAD	$C_{27}H_{33}O_{15}N_9P_2$	Flavine adeninen dinucleotid
FADH2	$C_{27}H_{35}O_{15}N_9P_2$	
FBP	$C_6H_{14}O_{12}P_2$	Fructose-1-6-biphosphate
FTHF	$C_{20}H_{23}O_7N_7$	Formyl tetrahydrofolate
Fum	$C_4H_4O_4$	Fumarate
G1P	$C_6H_{13}O_9P$	Glucose-1-phosphate
G3P	$C_3H_7O_6P$	Glyceraldehyde-3-phosphate
G6P	$C_6H_{13}O_9P$	Glucose-6-phosphate
GA1P	$C_6H_{14}O_8NP$	D-glucosamine-6-phosphate
GA6P	$C_6H_{14}O_8NP$	D-glucosamine-6-phosphate
GDP	$C_{10}H_{15}O_{11}N_5P_2$	Guanosine diphosphate
GLC	$C_6H_{12}O_6$	Glucose
Glcg	$C_6H_{10}O_5$	Glycogen
Gln	$C_5H_{10}O_3N_2$	Glutamine
Glu	$C_5H_9O_4N$	Glutamate
Gly	$C_2H_5O_2N$	Glycine
GMP	$C_{10}H_{14}O_8N_5P$	Guanosine monophosphate
Go3P	$C_3H_9O_6P$	Glycerol-3-phosphate
GTP	$C_{10}H_{16}O_{14}N_5P_3$	Guanosine triphosphate
H	H^+	Hydrogene
H2CO3	CH_2O_3	Bicarbonate
H2O	H_2O	Water
H2O2	H_2O_2	
H2S	H_2S	Hydrogene sulfide
H2SO4	H_2O_4S	Sulfuric acid
His	$C_6H_9O_2N_3$	Histidine
HSer	$C_4H_9O_3N$	Homoserine
iCit	$C_6H_8O_7$	isocitraat
Ile	$C_6H_{13}O_2N$	Isoleucine
IMP	$C_{10}H_{13}O_8N_4P$	Inosine monophosphate
Leu	$C_6H_{13}O_2N$	Leucine
Lipa	$C_{110}H_{196}O_{32}N_2P_2$	Lipid A
Lipid	$C_{40.2}H_{77.6}O_{8.41}N_{0.771}P_{1.03}$	Lipid composition
Lps	$C_{171}H_{298}O_{81}N_4P_2$	Lipo Poly sacharide
Lys	$C_6H_{14}O_2N_2$	Lysine
Mal	$C_4H_6O_5$	Malate
MalACP	$C_3H_3O_3Pept$	Malonyl ACP
MalCoA	$C_{24}H_{34}O_{19}N_7P_3S$	Malonyl CoA
MDAP	$C_7H_{14}O_4N_2$	Meso-diaminopimelate
Met	$C_5H_{11}O_2NS$	Methionine
MeTHF	$C_{20}H_{23}O_6N_7$	Methyleen tetrahydro folate
MTHF	$C_{20}H_{25}O_6N_7$	Methyl tetrahydrofolate
NAD	$C_{21}H_{28}O_{14}N_7P_2^+$	Nicotinamide adenine dinucleotide
NADH	$C_{21}H_{29}O_{14}N_7P_2$	
NADP	$C_{21}H_{28}O_{17}N_7P_3^+$	Nicotinamide adenine dinucleotide phosphate
NADPH	$C_{21}H_{29}O_{17}N_7P_3$	
NH3	H_3N	Ammonia
O2	O_2	Oxygen
OAA	$C_4H_4O_5$	Oxaloacetate
Orn	$C_5H_{12}O_2N_2$	Ornithine
PA	$C_{37}H_{71}O_8P$	Phosphatidyl acid
PAP	$C_{10}H_{15}O_{10}N_5P_2$	Phospho adenosine phosphate
PEP	$C_3H_5O_6P$	Phosphoenolpyruvate
Peptido	$C_{35}H_{53}O_{16}N_7$	Peptidoglycane
PEthAn	$C_{39}H_{76}O_8NP$	Phosphatidyl ethanolamine
PG	$C_{40}H_{75}O_9P$	Phosphatidyl glycerol
Phe	$C_9H_{11}O_2N$	Phenylalanine
PiOH	H_3O_4P	Phosphate
PPiOH	$H_4O_7P_2$	Pyrophosphate
Pro	$C_5H_9O_2N$	Proline
Prot	$C_{4.8}H_{9.67}O_{2.4}N_{1.37}S_{0.046}$	Protein composition
PRPP	$C_5H_{13}O_{14}P_3$	5-phospho-alpha-D-ribosyl-1-pyrophosphate
PSer	$C_{40}H_{76}O_{10}NP$	Phosphatidyl Serine
Pyr	$C_3H_4O_3$	Pyruvate
Qa	$C_7H_{12}O_6$	Quinate
R5P	$C_5H_{11}O_8P$	Ribose-5-phosphate
Rl5P	$C_5H_{11}O_8P$	Ribulose-5-phosphate
RNA	$C_{9.58}H_{13.8}O_{7.95}N_{3.95}P$	RNA composition
S7P	$C_7H_{15}O_{10}P$	Sedoheptulose-7-phosphate
Ser	$C_3H_7O_3N$	Serine
Shi	$C_7H_{10}O_5$	Shikimate
Shi3P	$C_7H_{11}O_8P$	Shikimate-3-phosphate
Suc	$C_4H_6O_4$	Succinate
SucCoA	$C_{25}H_{36}O_{19}N_7P_3S$	Succinyl CoA
THF	$C_{19}H_{23}O_6N_7$	Tetrahydrofolate
Thiored	Pept	Thioredoxin
ThioredH2	H_2Pept	Reduced thioredoxin
Thr	$C_4H_9O_3N$	Threonine
Trp	$C_{11}H_{12}O_2N_2$	Tryptophan
Tyr	$C_9H_{11}O_3N$	Tyrosine
UDP	$C_9H_{14}O_{12}N_2P_2$	Uridine diphosphate
UDPGlc	$C_{15}H_{24}O_{17}N_2P_2$	UDP glucose
UDPNAG	$C_{17}H_{27}O_{17}N_3P_2$	UDP N-acetyl glucosamine
UMP	$C_9H_{13}O_9N_2P$	Uridine monophosphate
UTP	$C_9H_{15}O_{15}N_2P_3$	Uridine triphosphate
Val	$C_5H_{11}O_2N$	Valine
XMP	$C_{10}H_{13}O_9N_4P$	Xanthosine-5-phosphate
Xu5P	$C_5H_{11}O_8P$	Xylulose-5-phosphate

Figure 13. Appendix B: List of metabolites. The first column contains the names as used in Appendix A. The second column contains the elemental composition and the charge: *Pept* stands for polypeptide, i.e., a complex molecule for which it makes no sense to use the exact elemental composition. The third column gives the full name of the metabolites.

5. Conclusions

With the aid of MFA (of which an overview was given, including the error propagation of the measured fluxes to the calculated ones and how to use redundant measurements to test for consistency of the data), it was shown that in the *E. coli* W3110.shik1 strain, the difference in shikimic acid yield between carbon-rich and carbon-limited cultures is not due to a lower flux in the aromatic amino acid pathway but to a larger excretion of dehydroshikimic acid and dehydroquinic acid. The flux entering the shikimate pathway was even higher in carbon-limited cultures, and this was due to or caused by less carbon going from the pentose phosphate pathway to the glycolysis.

Flux data were compared with RNA expression data. Most of the fluxes were not driven by the amount of expression of the corresponding gene. However, some fluxes were correlated, especially those from the altered genes: *aroF* and *aroG* feedback inhibition, giving a higher flux through the aromatic pathway, and *ydiB* overexpression, causing a high quinate production. Finally, a number of important fluxes are negatively correlated with the RNA expression level: this to maximize the flux through that reaction, for example, *ptsG* under carbon-limitation.

Acknowledgment

We thank Dr. Walter van Gulik of the Technical University Delft (TUDelft) for taking care of the elemental analysis of the biomass. This project was funded by the FWO-Vlaanderen (Fund for Scientific Research, Flanders, Belgium) grant G.0184.04 and by a Marie-Curie scholarship (training site: QLK-3-CT-2001-60077) from the EU fifth framework program. Jo Maertens is a research assistant of the FWO-Vlaanderen.

References and Notes

(1) De Clercq, E. *Nat. Rev./Drug Discovery* **2002**, *1*, 13−25.
(2) Johansson, L.; Lindskog, A.; Silfversparre, G.; Cimander, C.; Nielsen, K. F.; Lidén, G. *Biotechnol. Bioeng.* **2005**, *92*, 541−552.
(3) Savinell, J.; Palsson, B. *J. Theor. Biol.* **1992**, *154*, 421−454.
(4) van Gulik, W. M.; de Laat, W. T. A. M.; Vinke, J. L.; Heijnen, J. J. *Biotechnol. Bioeng.* **2000**, *68*, 602−618.
(5) Wang, N.; Stephanopoulos, G. *Biotechnol. Bioeng.* **1983**, *25*, 2177−2208.
(6) van der Heijden, R.; Romein, B.; Heijnen, J.; Hellinga, C.; Luyben, K. *Biotechnol. Bioeng.* **1994**, *43*, 11−20.
(7) Romein, B. Mathematical Modelling of Mammalian Cells in Suspension Culture. Ph.D. Thesis, Technische Universiteit Delft, The Netherlands, 2001.
(8) Stephanopoulos, G.; Aristidou, A.; Nielsen, J. *Metabolic Engineering. Principles and Methodologies*; Academic Press: New York, 1998.
(9) Nielsen, J.; Villadsen, J.; Lidén, G. *Bioreaction Engineering Principles*, 2nd ed.; Kluwer Academic/Plenum Publishers: New York, 2003.
(10) Knop, D. R.; Draths, K. M.; Chandran, S. S.; Barker, J. L.; von Daeniken, R.; Weber, W.; Frost, J. W. *J. Am. Chem. Soc.* **2001**, *123*, 10173−10182.
(11) Huang, G. S.; Hong, M.-Y.; Liu, Y.-C. *Life Sci.* **2003**, *72*, 2525−2531.
(12) Åkesson, M.; Förster, J.; Nielsen, J. *Metab. Eng.* **2004**, *6*, 285−293.
(13) Battley, E. H. *Appl. Environ. Microbiol.* **1995**, *61*, 1655−1657.
(14) Long, A.; Mangalam, H.; Chan, B.; Tolleri, L.; Hatfield, G.; Baldi, P. *J. Biol. Chem.* **2001**, *276*, 19937−19944.
(15) Johansson, L.; Lidén, G. *J. Biotechnol.* **2006**, in press.
(16) van der Heijden, R.; Heijnen, J. Unpublished work, 1995.
(17) Noorman, H.; Luyben, K.; Heijnen, J. *Biotechnol. Bioeng.* **1996**, *49*, 364−376.
(18) Golub, G.; Loan, C. V. *Matrix Computations*, 3rd ed.; Johns Hopkins University Press: Baltimore and London, 1996.
(19) van der Heijden, R. State Estimation and Error Diagnosis for Biotechnological Processes. Ph.D. Thesis, Technische Universiteit Delft, Nederland, 1991.
(20) Lequeux, G.; Van der Heijden, R.; Van Den Broeck, S.; Vanrolleghem, P. In *Proceedings of the 9th IFAC Conference on Computer Applications in Biotechnology CAB9*, Nancy, France, 2004.
(21) Schilling, C.; Schuster, S.; Palsson, B.; Heinrich, R. *Biotechnol. Prog.* **1999**, *15*, 296−303.
(22) van der Heijden, R.; Heijnen, J.; Hellinga, C.; Romein, B.; Luyben, K. *Biotechnol. Bioeng.* **1994**, *43*, 3−10.
(23) Madron, F.; Veverka, V.; Vanecek, V. *AIChE J.* **1977**, *23*, 482−486.
(24) Reilly, P.; Carpani, R. In *13th Chemical Engineering Conference*; Montreal, Canada, 1963.
(25) Carlson, R.; Srienc, F. *Biotechnol. Bioeng.* **2004**, *85*, 1−19.
(26) Pramanik, J.; Keasling, J. *Biotechnol. Bioeng.* **1997**, *56*, 398−421.
(27) Chen, R.; Yap, W. M. G. J.; Postma, P. W.; Bailey, J. E. *Biotechnol. Bioeng.* **1997**, *56*, 583−590.
(28) Noronha, S. B.; Yeh, H. J. C.; Spande, T. F.; Shiloach, J. *Biotechnol. Bioeng.* **2000**, *68*, 316−327.
(29) Phue, J.-N.; Shiloach, J. *J. Biotechnol.* **2004**, *109*, 21−30.
(30) Cozzone, A. J. *Ann. Rev. Microbiol.* **1998**, *52*, 127−164.
(31) Sauer, U.; Lasko, D. R.; Fiaux, J.; Hochuli, M.; Glaser, R.; Szyperski, T.; Wüthrich, K.; Bailey, J. E. *J. Biotechnol.* **1999**, *181*, 6679−6688.
(32) Wick, L. M.; Quadroni, M.; Egli, T. *Environ. Microbiol.* **2001**, *3*, 588−599.
(33) Draths, K.; Frost, J. *J. Am. Chem. Soc.* **1991**, *113*, 9361−9363.
(34) Keseler, I. M.; Collado-Vides, J.; Gama-Castro, S.; Ingraham, J.; Paley, S.; Paulsen, I. T.; Peralta-Gil, M.; Karp, P. D. *Nucleic Acids Res.* **2005**, *33*, D334−D337.
(35) Reed, J. L.; Vo, T. D.; Schilling, C. H.; Palsson, B. *Genome Biol.* **2003**, *4*, R54.
(36) Kanehisa, M.; Goto, S. *Nucleic Acids Res.* **2000**, *28*, 27−30.
(37) Varma, A.; Palsson, B. *J. Theor. Biol.* **1993**, *165*, 477−502.
(38) Draths, K. M.; Pompliano, D. L.; Conley, D. L.; Frost, J. W.; Berry, A.; Disbrow, G. L.; Staversky, R. J.; Lievense, J. C. *J. Am. Chem. Soc.* **1992**, *114*, 3956−3962.
(39) Varma, A.; Palsson, B. *Bio/Technology* **1994**, *12*, 994−998.
(40) Kayser, A.; Weber, J.; Hecht, V.; Rinas, U. *Microbiology* **2005**, *151*, 693−706.
(41) Carlson, R.; Srienc, F. *Biotechnol. Bioeng.* **2004**, *86*, 149−162.
(42) Oh, M.-K.; Liao, J. C. *Biotechnol. Prog.* **2000**, *16*, 278−286.
(43) Hua, Q.; Yang, C.; Oshima, T.; Mori, H.; Shimizu, K. *Appl. Environ. Microbiol.* **2004**, *70*, 2354−2366.
(44) Shin, S.; Song, S. G.; Lee, D. S.; Pan, J. G.; Park, C. *FEMS Microbiol. Lett.* **1997**, *146*, 103−108.

Accepted for publication May 2, 2006.

BP060014U