Computational Design of Reduced Metabolic Networks

Scott Holzhütter[b] and Hermann-Georg Holzhütter*[a]

Cellular functions are based on thousands of chemical reactions and transport processes, most of them being catalysed and regulated by specific proteins. Systematic gene knockouts have provided evidence that this complex reaction network possesses considerable redundancy, that is, alternative routes exist along which signals and metabolic fluxes may be directed to accomplish an identical output behaviour. This property is of particular importance in cases where parts of the reaction network are transiently or permanently impaired, for example, due to an infection or genetic alterations. Here we present a computational concept to determine enzyme-reduced metabolic networks that are still sufficient to accomplish a given set of cellular functions. Our approach consists of defining an objective function that expresses the compromise that has to be made between successive reduction of the network by omission of enzymes and its decreasing thermodynamic and kinetic feasibility. Optimisation of this objective function results in a linear mixed-integer program. With increasing weight given to the reduction of the number of enzymes, the total flux in the network increases and some of the reactions have to proceed in thermodynamically unfavourable directions. The approach was applied to two metabolic schemes: the energy and redox metabolism of red blood cells and the carbon metabolism of Methylobacterium extorquens*. For these two example networks, we determined various variants of reduced networks differing in the number and types of disabled enzymes and disconnected reactions. Using a comprehensive kinetic model of the erythrocyte metabolism, we assess the kinetic feasibility of enzyme-reduced subnetworks. The number of enzymes predicted to be indispensable amounts to 14 (out of 28) for the erythrocyte scheme and 13 (out of 77) for the bacterium scheme, the largest group of enzymes predicted to be simultaneously dispensable amounts to 3 and 37 for these two systems. Our approach might contribute to identifying potential target enzymes for rational drug design, to rationalising gene-expression profiles of metabolic enzymes and to designing synthetic networks with highly specialised metabolic functions.*

Introduction

Complex cellular functions, such as motility, defence against toxic compounds, repair and replacement of impaired macromolecules and excretion of signalling compounds for communication with other cells, ultimately rest upon a complex network of well-coordinated chemical reactions and transport processes, most of them being catalysed by enzymes and facilitated by transport proteins. Throughout this paper the terms "reactions" and "enzymes" will be used to designate both chemical reactions and spatial transport processes as well as any type of protein directly controlling the activity of a reaction. Enzyme activities can be modulated by various modes of regulation such as allosteric effectors, reversible phosphorylation and variable gene expression. These regulatory mechanisms have evolved naturally and enable the cell to activate and inactivate parts of the total reaction network according to temporally varying functional requirements.

The synthesis of an enzyme is connected with the consumption of energy and external resources (e.g., essential amino acids). Thus, an economically reasonable way of activating functionally relevant subnetworks and inactivating temporarily irrelevant ones should consist of an effective regulation of enzyme synthesis and degradation. Time-dependent gene expression enables the cell to follow such a plausible strategy. For most cell proteins, the rate of their synthesis is controlled by transcriptional activity, that is, the number of related mRNA copies available at the ribosomes. Microarray techniques have made it possible to monitor the mRNA levels of thousands of different proteins in a time-dependent manner. Analysis of such gene-expression profiles have provided evidence that complete metabolic pathways can be switched on or off depending on their relative importance for the accomplishment of a specific pattern of cellular functions. For example, yeast cells respond to progressive glucose depletion by turning down the expression of glycolytic enzymes whilst turning on the expression of enzymes belonging to those pathways that are needed for an efficient utilisation of ethanol serving as a substitute for glucose.[1] Similar gene-expression studies have been performed by Korke et al., who compared the gene-expression profiles of mammalian cells grown in culture under low and high molar ratios of glucose consumption to lactate production (0.08 and 1.4, respectively).[2] They found 1.4–3.0

[a] *Prof. Dr. H.-G. Holzhütter*
Humboldt-University Berlin, Medical Department (Charité)
Institute of Biochemistry
Monbijoustraße 2, 10117 Berlin (Germany)
Fax: (+49) 30-450-528-937
E-mail: hergo@charite.de

[b] *S. Holzhütter*
Technical University Berlin, Institute of Mathematics
Straße des 17. Juni 135, 10623 Berlin (Germany)

 DOI: 10.1002/cbic.200400128

fold changes in the expression level of most enzymes involved in glucose metabolism. These changes were accompanied by three- to fivefold changes of the related metabolic flux rates. Remarkably, most of the glycolytic enzymes identified as differentially expressed catalyse unregulated, reversible reactions. As the reduction of the glycolytic flux can be achieved by altering the level of phosphofructokinase alone, such a concerted down-regulation of the whole pathway points to a remarkable economisation of enzyme usage. A further example demonstrating the capability of cells to express genes of metabolic enzymes "just in time" for achieving a maximal metabolic output comes from recent work by Zaslaver et al.[3] They observed that the promoters of various enzymes involved in amino acid biosynthesis pathways of *Escherichia coli* are consecutively activated according the enzyme order along the pathway. Such a wave-like activation of enzymes has been predicted in ref. [4] as an optimal strategy to rapidly reach a production goal at minimal total enzyme production.

These observations raise the general question of which enzymes and associated reactions of the complete cellular reaction network can be simultaneously down-regulated without compromising those "target" reactions indispensable for accomplishing a well-defined functional state of the cell. To address this problem, we have developed a mathematical approach that allows smaller subnetworks comprising a reduced number of enzymes but still capable of producing a required cellular output to be identified within the complete reaction network. The decision as to whether or not any reduced set of enzymes is indeed sufficient to maintain a stationary flux regime at given side constraints is complicated by the fact that the possible magnitude and direction of metabolic fluxes is determined by kinetic and thermodynamic constraints that are not, or only fragmentarily, known for most reactions.

Kinetic constraints on the maximal flux through an enzyme-catalysed reaction arise from the maximal catalytic capacity (= turnover number) of the enzyme and the enzyme concentration that can be realistically achieved under cellular conditions. Even if these two parameters are known, they only provide a rough estimate of the maximal possible flux as the catalytic capacity of an enzyme can be drastically reduced due to various modes of inhibition, such as allosteric inhibitors or enzyme phosphorylation. The concentration of an enzyme can also be very low due to a low transcription rate. Thus, putting kinetically realistic upper bounds to the magnitude of the metabolic fluxes requires a profound knowledge of the in vivo kinetics and genetic control of the underlying enzymes. This knowledge is currently only available for a very limited number of metabolic pathways, such as the main metabolic pathways of erythrocytes[5–11] or glycolysis in yeast cells.[12] Thus, to make our approach applicable to arbitrary metabolic networks with unknown in vivo enzyme kinetics but known stoichiometry we have adopted the well-established concept of flux-balance analysis (FBA)[13, 14] and extended it by introducing the costs for enzyme synthesis and the thermodynamic feasibility of fluxes as additional network properties that may be subjected to optimisation. FBA makes use of the fact that under steady-state conditions the sum of fluxes producing or degrading any "internal" metabolite has to be zero. Application of this method is based on only two prerequisites: 1) the topology of the metabolic network under consideration has to be known and 2) an evaluation criterion is needed to pick out the most likely flux distribution among all those flux distributions that are compatible with the steady-state conditions. The topology of the metabolic network is given in terms of the so-called stoichiometric matrix, which relates the time-dependent variation of the metabolite concentrations to the fluxes through all metabolic processes for which an enzyme or transport protein is available in a given cell type. The topology of central metabolic pathways is meanwhile available for numerous cell types (see for example, http://www.genome.ad.jp/kegg).

In previous applications of FBA, the maximal production of biomass at restraint influx of substrates has been used as evaluation criterion.[15, 16] However, maximisation of biomass production makes little sense for higher eukaryotic cells with multiple functions. Therefore, the minimisation of the weighted sum of the internal fluxes has recently been proposed as a novel and more general flux-evaluation criterion[17] whereby the thermodynamic equilibrium constants are used as weighting factors, that is, the more the thermodynamic equilibrium lies on the right-hand side of the reaction, the larger the weighting factor (= costs) for the backward reaction. This differential weighting of forward and backward fluxes takes into account thermodynamic constraints of the flux directions arising from the Gibb's free energy of the reactions. The net reaction always proceeds in that direction which is associated with a decrease of Gibb's free energy (ΔG). The value of ΔG, however, depends upon the concentrations of the reactants, which may change by several orders of magnitude when switching the metabolic network from one functional state to another, for example, from a glucose-consuming to a glucose-producing state.[18] This makes it difficult to put rigid constraints on the directionality of the fluxes. Instead, our approach will take into account the value of the free-energy change at standard conditions (ΔG_0) by an appropriate weighting of the fluxes. The larger the absolute value of ΔG_0, the greater the work that has to be expended by the network to generate concentration gradients capable of driving a reaction into a direction opposed to the direction dictated by ΔG_0.

The core of our approach is a mathematical objective function that evaluates the "costs" accruing for the cell on one hand for the synthesis of enzymes and on the other hand for the accomplishment of a balanced flux distribution. Identification of subnetworks consists of the minimisation of this objective function with increasing weight given to the costs of enzyme synthesis. The objective function will be constructed as a linear function with respect to both the flux rates and the Boolean variables designating whether an enzyme is present or not. This linearity in the variables allows the objective function to be minimised by means of the well-elaborated mathematical apparatus of mixed-integer programming.[19] In the second part of this article we demonstrate the application of our method to two metabolic schemes of different complexity: the energy and redox metabolism of red blood cells and the carbon metabolism of *Methylobacterium extorquens*.

Computational Methods

Definition of basic quantities

We define the complete metabolic network under consideration by the constituting reactions R_j ($j=1,2,\ldots,n_r$) forming the reaction vector **R**, the flux rates v_j through these reactions forming the flux vector **V**, the enzymes E_k ($k=1,2,\ldots,n_e$) capable of catalysing at least one of the reactions forming the enzyme vector **E** and the metabolites S_i ($i=1,2,\ldots,n_s$) involved in the reactions forming the metabolite vector **S**. To specify the relations between reactions, enzymes and metabolites we introduce the following entities.

The elements $N_{i,j}$ of the stoichiometric matrix **N** indicate how flux v_j through reaction R_j affects the concentration of metabolite S_i: $N_{i,j}>0 \rightarrow N_{i,j}$ molecules of metabolite i are formed during a single reaction, $N_{ij}<0 \rightarrow N_{i,j}$ molecules of metabolite i are consumed during a single reaction j, $N_{ij}=0 \rightarrow$metabolite i is not involved in reaction j. For example, for the flux v_8 through the chemical reaction

$$2\,S_1 + S_2 \xrightarrow{v_8} S_3 + 3\,S_4$$

the elements of the stoichiometric matrix read $N_{18}=-2$, $N_{28}=-1$, $N_{38}=1$, $N_{48}=3$.

To relate the reactions of the network to the participating enzymes, we introduce reaction-specific enzyme sets L_j containing all indices of enzymes capable of catalysing reaction R_j. For the sake of completeness, we will introduce pseudo-enzymes for those reactions or transport processes proceeding spontaneously (e.g. formation of reactive oxygen species or free membrane diffusion). By definition, the reaction-specific enzyme set L_j must contain at least one enzyme index, otherwise the reaction can be neglected. If L_j contains more than one index, that is, if reaction R_j can be catalysed by several enzymes, it seems reasonable to classify these enzymes according to their relative capacities. Enzyme E_k is a so-called *master enzyme* for reaction R_j if it catalyses this reaction with significantly higher efficiency than all other reactions of the network. Enzyme E_k is a so-called *assistant enzyme* for reaction R_j if it catalyses the reaction with significantly lower efficiency than at least one other reaction. For example, in hepatocytes, the phosphorylation of glucose to glucose-6-phosphate is catalysed by the unspecific hexokinase as well as by the liver-specific glucokinase whereby the glucokinase possesses a much higher activity than the hexokinase and thus represents the master enzyme for this reaction.

The expression status of an arbitrary enzyme E_k is described by the binary decision variable d_k, which may assume the values 1 or 0 to indicate that the enzyme is either available or not. The decision variables form the expression vector **D**. If the L_1-norm of **D** is smaller than the total number of enzymes

$$\| \mathbf{D} \| = \sum_{k=1}^{n_e} d_k < n_e$$

we will speak of an enzyme-reduced subnetwork.

In general, the fluxes v_j may be positive or negative, that is, the net reaction may proceed either in the forward or backward direction. To deal with non-negative variables, we write the flux as the difference of the two positive-definite fluxes $v_j^{(+)}$ and $v_j^{(-)}$, which represent the absolute magnitude of the flux if the net reaction proceeds either in the forward or backward direction:

$$v_j = v_j^{(+)} - v_j^{(-)} \qquad (1)$$

$$v_j^{(+)} = v_j\,\Theta(v_j),\ v_j^{(-)} = v_j\,[\Theta(v_j)-1]$$

$\Theta(x)$ denotes the unit-step function, that is, by definition only one of the two components $v_j^{(+)}$ and $v_j^{(-)}$ can be different from zero at any one time. This condition should not be used as a side constraint since it is always satisfied due to flux minimisation. The forward direction is defined as that which would ensure a positive Gibbs free-energy change under standard conditions (where all reagents are present at unit concentrations); at these standard conditions the backward flux is defined as zero.

As pointed out above, a single reaction can be catalysed by several enzymes, that is, the reaction-specific enzyme set of a reaction may contain more than one enzyme index. Accordingly, the flux through a reaction can be split into the partial fluxes through the various enzymes capable of catalysing this reaction:

$$v_j = \sum_{k\in L_j} w_{j,k} = \sum_{k\in L_j} \left(w_{j,k}^{(+)} - w_{j,k}^{(-)}\right) \qquad (2)$$

Here $w_{j,k}^{(+)}$ and $w_{j,k}^{(-)}$ ($j=1,2,\ldots,n_r$,) denote the absolute magnitude of the partial forward and backward flux through reaction R_j catalysed by enzyme E_k. Note that the number of partial fluxes for reaction R_j equals the number of elements of the reaction-specific enzyme set L_j.

Evidently, the partial fluxes $w_{j,k}^{(+/-)}$ can only be different from zero if enzyme E_k is present, that is, $d_k>0$. This condition can be formulated by the inequality relation:

$$n_r\,W_{\max}\,d_k \geq \sum_{j=1}^{n_r} \left(w_{j,k}^{(+)} + w_{j,k}^{(-)}\right) \qquad (3)$$

here the positive constant $n_r W_{\max}$ on the left-hand side represents the product between the total number of reactions n_r and the highest possible flux $W_{\max}$ (see below).

Flux-balance conditions

Under steady-state conditions, the fluxes have to obey the flux-balance conditions:

$$\sum_{j=1}^{n_r} N_{ij}\,v_{ij} = \sum_{j=1}^{n_r} N_{ij}\left(v_j^{(+)} - v_j^{(-)}\right) = \sum_{j=1}^{n_r}\sum_{k=1}^{n_r} \left(w_{j,k}^{(+)} - w_{j,k}^{(-)}\right) = 0 \quad (i = 1,\ldots,n_s) \qquad (4)$$

Equations (4) constitute a homogeneous system with respect to the unknown fluxes and represent the principle of conservation of mass for a homogeneous reaction system. Depending on the choice of the enzyme decision variables d_k ($k=1,2,\ldots,n_e$) some of the partial fluxes have to be zero according to Equation (3) so that system (4) may possess either only the trivial solution $v_j^{(+)}=v_j^{(-)}=0$ (for all j) or an infinite number of non-zero solutions, as all equations in (4) can be multiplied by an arbitrary non-zero constant. In the latter case, additional criteria are needed to select a unique solution that is the most reliable one from the biochemical view point.

Fixing the target fluxes through functionally essential reactions

All cellular functions are ultimately linked to a certain number (n_t) of target reactions producing the metabolites that are required for building cellular structures and for driving functionally important reactions. For example, movement of sperm cells requires the synthesis of microtubules forming the flagella at the outer cell surface (=functional structure) and the synthesis of ATP yielding the energy for the flagella movement (=functional reaction). In general, defining the functions that the cell has to fulfil means to put the fluxes through a certain number (r_t) of target reactions to non-zero values:

$$v_j = T_j,\ T_j >= \ (j = j_1, j_2, \ldots j_{n_t}) \tag{5}$$

Some of the target reactions, such as the production of energy (ATP) or the synthesis of membrane phospholipids, are permanently required to ensure cell integrity. Other target reactions, for example the synthesis of a hormone or the detoxification of a pharmaceutical, may be only temporarily required.

Maximal fluxes

Due to physical laws governing the maximal number of collisions between an enzyme and its substrates per unit time, the flux capacity of any enzyme-catalysed reaction must have an upper boundary:

$$W_{j,k} < W_{\max} \tag{6}$$

The turnover rate of a catalytically perfect enzyme is determined by the on-rate constant for the formation of the enzyme–substrate complex, which is of the order $\approx 10^8\,\mathrm{M^{-1}\,s^{-1}}$.[20] The concentrations of substrates and enzymes are typically below 10^{-3} and 10^{-9} M, respectively. Thus, an ample estimate of the value of $W_{\max}$ should be $3.6\times10^2\,\mathrm{mM\,h^{-1}}$. This value was used in the following calculations.

Flux constraints arising from the availability of external metabolites

The non-equilibrium state of biochemical reaction systems is maintained by a steady uptake of energy-rich, low-entropy substrates and the release of low-energy, high-entropy products. The absence of a certain substrate associated with the exchange flux v_i can be expressed by forcing the uptake component of the flux to zero:

$$v_i^{(\mathrm{uptake})} = 0 \tag{7}$$

Thermodynamic constraints: Irreversibility of reactions

The direction of any flux v_j is dictated by the change of Gibbs free energy:

$$\Delta G_j = \Delta G_j^{(0)} + RT \ln\left(\frac{\prod_{i=1}^{n_s}[S_i]^{N_{ij}^{(+)}}}{\prod_{i=1}^{n_s}[S_i]^{N_{ij}^{(-)}}}\right) \tag{8}$$

with $N_{ij}^{(+)} = N_{ij}$ if $N_{ij} \geq 0$, $N_{ij}^{(-)} = -N_{ij}$ if $N_{ij} < 0$

$\Delta G_j^{(0)}$ denotes the change of free energy under the condition that all reagents are present at unit concentrations ($=1\,\mathrm{mol\,L^{-1}}$). $\Delta G_j^{(0)}$ can be expressed through the thermodynamic equilibrium constant K_j^{equ} by:

$$\Delta G_j^{(0)} = -RT \ln\left(K_j^{\mathrm{equ}}\right) \tag{9}$$

here $RT = 2.48\,\mathrm{kJ\,mol^{-1}}$ at room temperature ($T=25\,^\circ\mathrm{C}$). As stated above, we will notate all reactions of the network such that under standard conditions it holds that $\Delta G_j^{(0)} \leq 0$ ($K_j^{\mathrm{equ}} \geq 1$) and thus ($v_j > 0$ ($v_j^{(-)} = 0$)). The second term in the right-hand side of Equation (8) depends upon the actual concentrations of the reactants, which, under cellular conditions, may strongly deviate from unit concentrations. With accumulating concentrations of the reaction products (appearing in the nominator) and/or vanishing concentrations of the reaction substrates (appearing in the denominator) the concentration-dependent term in Equation (8) may assume arbitrarily large negative values, that is, in principle the direction of a chemical reaction can always be reversed provided that other reactions in the system are capable of accomplishing the required change in the concentration of the reactants. For example, the standard free-energy change of the glycolytic reaction (glycerol aldehyde phosphate→dihydroxy acetone phosphate) catalysed by the enzyme triose phosphate isomerase amounts to $=-7.94\,\mathrm{kJ\,mol^{-1}}$ (=24.6). Nevertheless, under cellular conditions this reaction proceeds backwards (dihydroxy acetone phosphate→glyceraldehyde phosphate), as the reaction substrate glycerol aldehyde phosphate is rapidly converted into 1,3-bisphosphoglycerate along the glycolytic pathway. This example shows that a sharp classification into reversible and irreversible reactions on the sole basis of $\Delta G^{(0)}$ can be problematic. Instead, we will use the value of the equilibrium constant as weighting factor in the definition of the flux-evaluation function (see below).

The flux-evaluation function

By definition, maintenance of cell integrity requires the target fluxes T_j in Equation (5) to be different from zero. This excludes system (4) from being satisfied by the trivial solution in which all fluxes do vanish. Depending on the choice of the expression vector **D** and of the constraints (3) and (5), equation system (4) may have i) no solution, ii) a unique solution (example: a linear chain of monomolecular reactions with given non-zero target flux for one of the reactions) or iii) an infinite number of solutions. In the latter case, we need an evaluation criterion to decide which of the possible solutions is the most reliable one. To this end, we define a flux-evaluation function that attains its minimum for the most reliable flux distribution. Based on the thermodynamic considerations above, the flux-evaluation function should punish those flux distributions in which reactions with large values of their equilibrium constant are forced to proceed in the reverse direction. This condition can be incorporated into the flux-evaluation function by weighting the backward fluxes more highly than the forward fluxes. For the choice of the weighting factor, we take into consideration the fact that the rate law of an enzyme-catalysed reaction has the general form:

$$v_j = v_{\max_j} F_j\left(\prod_{i=1}^{n_s}[S_i]^{N_{ij}^{(-)}} - \frac{1}{K_j^{\mathrm{equ}}}\prod_{i=1}^{n_s}[S_i]^{N_{ij}^{(+)}}\right) \tag{10}$$

Here $v_{\max_j}$ denotes the maximal rate of the enzyme with respect to reaction j, F_j (<1) is a nonlinear function with respect to the me-

tabolites S_i that captures specific kinetic properties of the enzyme, for example, binding of the reactants and of allosteric effectors. From this general form of the rate equation, it follows that to obtain the same magnitude of the flux in either the forward or backward direction requires the stoichiometric product $\prod_{i=1}^{n_s} [S]_i^{N_{ij}^{(+)}}$ of the reaction products to be K_j^{equ}-fold higher than the stoichiometric product of the substrates. We thus will choose the equilibrium constant as weighting factor for the backward fluxes.

A second aspect relevant to the definition of the flux-evaluation function pertains to multifunctional enzymes, that is, enzymes catalysing more than one reaction. Evidently, establishing a certain flux value through a reaction should be easier by using a master enzyme instead of using non-specialised assistant enzymes (for the distinction between these two categories of enzymes see the definitions above). Such a constraint can be incorporated into the flux-evaluation function by introducing the weights:

$$\beta_{j,k} = \begin{cases} \beta > 1 \text{ if enzyme } E_k \text{ is assistant enzyme for reaction } R_j \\ 1 \text{ if enzyme } E_k \text{ is master enzyme for reaction } R_j \end{cases} \quad (11)$$

of the enzyme-specific partial fluxes. Based on these arguments, we define the flux-evaluation function as follows:

$$\Theta_v(W^{(+)},W^{(-)}) = \sum_{j=1}^{n_r} \sum_{k \in L_j} \beta_{j,k} \left(\frac{1}{\sqrt{1+(K_j^{equ})^2}} W_{j,k}^{(+)} + \frac{K_j^{equ}}{\sqrt{1+(K_j^{equ})^2}} W_{j,k}^{(-)} \right) \quad (12)$$

We weight the partial forward and backward fluxes by the equilibrium constant in a reciprocal manner whereby the sum of squared weights is normalised to unity. Reciprocal weighting of the forward and backward fluxes is necessary because the value of the flux-evaluation function should not depend on the way we define the forward direction—changing this definition, that is, exchanging substrates and products, means inverting the equilibrium constant.

Evaluating the costs of a metabolic network

Let λ denote the average costs in terms of energy and external metabolites that have to be paid by the cell for keeping the concentration of a single enzyme on a level that is sufficient to accomplish a "typical" flux through the reaction catalysed by this enzyme. Accordingly, the total costs Θ_E for the realisation of an enzyme-expression pattern defined by the expression vector **D** are given by:

$$\Theta_E(\mathbf{D}) = \lambda \sum_{k=1}^{n_e} d_k \quad (13)$$

The value of the flux-evaluation function (12) can also be interpreted as the costs of maintaining the stationary flux distribution defined by the partial fluxes $w_{j,k}$. Thus, we measure the total costs for any subnetwork by the objective function:

$$\Theta(W^{(+)},W^{(-)},\mathbf{D}) = \Theta_v(W^{(+)},W^{(-)}) + \Theta_E(\mathbf{D}) = \sum_{j=1}^{n_r} \sum_{k=1}^{n_e} \beta_{j,k} \left(\frac{1}{\sqrt{1+(K_j^{equ})^2}} W_{j,k}^{(+)} + \frac{K_j^{equ}}{\sqrt{1+(K_j^{equ})^2}} W_{j,k}^{(-)} \right) + \lambda \sum_{k=1}^{n_e} d_k \quad (14)$$

The minimum of the objective function (14) defines the enzyme-expression status (expressed or non-expressed) and the stationary flux distribution of the network at a given value of the cost factor λ and imposed constraints (3), (4), (5), (6) and (7).

For the extreme case that $\lambda = 0$, that is, if there is no pressure on the costs for enzyme expression, the minimum of the objective function defines the so-called *flux-minimised* metabolic state, in which the flux-evaluation function assumes its absolute minimum. The flux-minimised metabolic state refers to a situation in which no restrictions exist that limit the cellular capacity to synthesise all enzymes of the network under study. Minimisation of the objective function (14) at $\lambda = 0$ is the mathematical formulation of the principal of flux minimisation proposed in a previous work.[17]

Increasing the value of the cost factor λ, that is, increasing the relative share of the costs for enzyme expression in the total costs of the network, the cell will be progressively forced to economise enzymes that are dispensable without compromising the establishment of a flux distribution that meets all side constraints. Hence, by increasing the value of λ we may generate subnetworks reduced in the number of enzymes and reactions. The minimum of the objective function attained if $\lambda \to \infty$ defines the *enzyme-minimised* subnetwork. It comprises the minimal number of enzymes required to generate a flux distribution still satisfying all constraints introduced above.

The problem is that the values of the cost factor λ at which the minimum of the objective function changes and defines a new subnetwork are not known a priori. Therefore, to generate all possible variants of reduced networks would require calculating the minimum of the objective function (14) on a sufficiently fine discrete grid of λ values. Even for the relatively small networks studied in this article, such a strategy turns out to be very expensive in terms of computing time. But even if this more technical problem could be resolved, the question remains whether the chosen discrete grid of λ values was fine enough to capture all interesting and biologically meaningful solutions. Because of this complication, we have decided to follow another strategy to generate subnetworks.

We put the decision variable of one enzyme after the other equal to zero and determine the minimum of the objective function (14) at a value of the cost factor λ that exceeds the maximum value of the flux-evaluation function, that is, we consider the n_e minimisation problems:

$$\mathrm{MIN}\{\Theta(W^{(+)},W^{(-)},\mathbf{D}) | \lambda = \lambda_0 | d_k = 0\} \quad (15)$$

with:

$$\lambda_0 = \underset{W^{(+)},W^{(-)}}{\mathrm{MAX}} \{\Theta_v(W^{(+)},W^{(-)})\} \quad (16)$$

Because of the general upper limit W_{max} imposed on all fluxes according to constraint (6), the maximal value of the flux-evaluation function can be estimated to:

$$\underset{W^{(+)},W^{(-)}}{\mathrm{MAX}} \{\Theta_v(W^{(+)},W^{(-)})\} < n_r \beta W_{max} \quad (17)$$

Hence condition (16) will be met by putting:

$$\lambda_0 = n_r \beta W_{max} \quad (18)$$

Solving the minimisation problems (15) for enzyme E_k (i.e. $d_k = 0$) defines the so-called *k-subnetwork*. It lacks all enzymes that are dis-

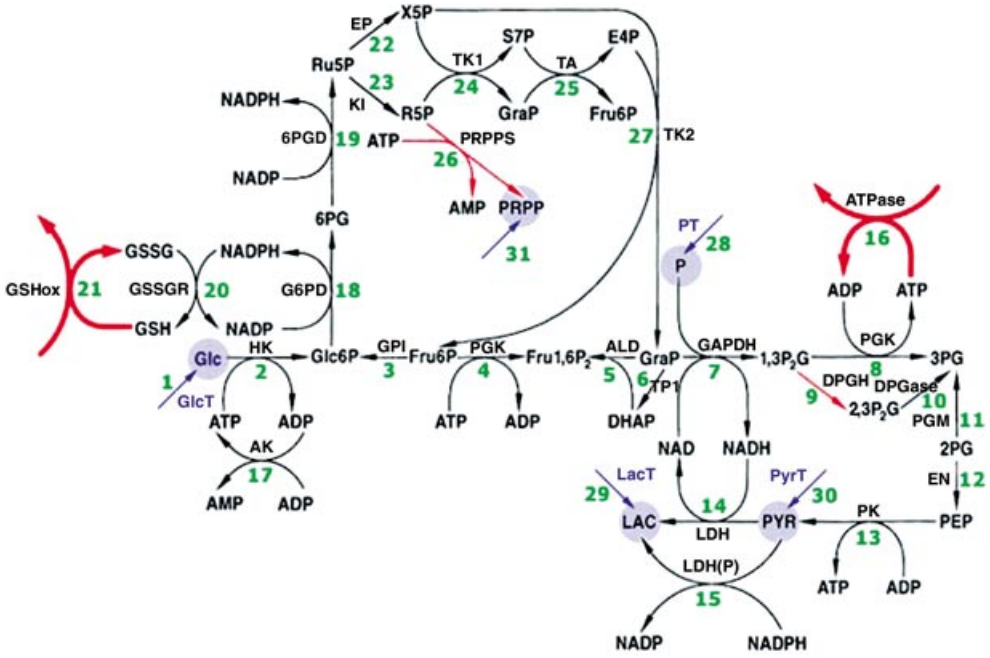
pensable under the prerequisite that enzyme E_k—the *generating enzyme*—is discarded. These dispensable enzymes together with the generating enzyme constitute the *k-enzyme cluster*. Accordingly, we define the *k-reaction cluster* as the group of all reactions of the *k*-subnetwork having zero flux ($v_j = 0$). If no solution of the minimisation problem (15) exists, the generating enzyme E_k is called *essential*, that is, it cannot be discarded without compromising the functionality of the network. Solving the minimisation problem (15) with respect to all enzymes ($k = 1,2,\ldots,n_e$), we may identify all essential enzymes and all variants of subnetworks characterised by the maximal number of enzymes that can be conjointly discarded.

Results

1. Application to a metabolic scheme of human red blood cells (erythrocytes)

As a first example for the application and reliability of our method, we have chosen the representative metabolic scheme of erythrocyte metabolism shown in Scheme 1. The scheme takes into account two cardinal metabolic pathways of this cell: glycolysis, including the so-called 2,3-bisphosphoglycerate (2,3-P_2G) shunt, and the pentose phosphate cycle, comprising an oxidative and a non-oxidative part. The network comprises 30 reactions, 28 enzymes and 29 metabolites of which only 25 metabolites are independent because four conservation conditions exist: AMP + ADP + ATP = const. = A (total adenine nucleotides), NAD + NADH = const. = ND, NADP + NADPH = const. = NDP and GSH + GSSG = const. = G. In Scheme 1, the orientation of the arrows corresponds to the "natural" (= forward) direction of the reactions which, as noted above, is the direction of the net reaction at standard conditions. Note, however, that in the optimisation calculations no a priori assumption is made about the directionality of fluxes, that is, all fluxes may be directed in both direction with higher costs for the backward direction.

The functionally essential target reactions that have to be maintained by the network are the following: i) formation of 2,3-P_2G (flux v_9) required to modulate the oxygen affinity of haemoglobin, ii) ATP-utilisation (flux v_{16}), which is mostly spent on the Na/K-ATPase to build up the Na^+/K^+ gradient across the plasma membrane, iii) oxidation of glutathione (GSH; flux v_{21}) to prevent oxidative damage of cellular proteins and lipids, iv) synthesis of phosphoribozyl diphosphate (PRPP; flux v_{26}) required to salvage adenine nucleotides. The magnitude of these four target reactions depends on the specific external conditions of the cell, for example, osmolarity of the blood or preservation medium, oxidative stress caused by reactive oxygen species or lowering of the oxygen tension during hypoxia. In our calculations, the flux values for these four target reactions were chosen as reported for the normal in vivo state of erythrocytes:[9] $T_9 = 0.49$, $T_{16} = 2.38$, $T_{21} = 0.093$, $T_{26} = 0.026$ mmol h^{-1}. Table 1 depicts the stoichiometric matrix and the reaction-specific enzyme sets associated with Scheme 1. The equilibrium constants for the reactions and the notation of the enzymes are given in Table 2. The equilibrium constants were taken from ref. [8]. Inspection of the link matrix reveals that most of the reactions are catalysed by a single enzyme, the master enzyme. Only reactions R_9, R_{10} and R_{11} can also be catalysed by an assistant enzyme. For example, for reaction R_9 (1,3P_2G → 2,3P_2G) enzyme E_9 (bisphosphoglycerate mutase) is the master enzyme, but this reaction can also be catalysed by the assistant enzyme E_{11} (phosphoglycerate mutase) although with



Scheme 1. *Reaction scheme of the metabolic network in erythrocytes analysed. Note that the reaction arrows point into the direction of the net reaction under standard conditions which for the reactions R_3, R_5, R_6, R_7, R_{11} and R_{29} differs from the direction under in vivo conditions. Reactions, enzymes, equilibrium constants and metabolites are explained in Tables 1 and 2.*

Table 1. *Stoichiometric matrix of the erythrocyte network.*[a]

			Fluxes																												
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
metabolites	1	Glcin	**1**	**−1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	Glc6P	0	**1**	**3**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**−1**	0	0	0	0	0	0	0	0	0	0	0	0
	3	Fru6P	0	0	**−1**	**−1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	**1**	0	0	0
	4	Fru16P2	0	0	0	**1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	GraP	0	0	0	0	**−1**	**−1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**−1**	0	**1**	0	0	0
	6	DHAP	0	0	0	0	**−1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	13P2Gri	0	0	0	0	0	0	**−1**	**−1**	**−1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	8	23P2Gri	0	0	0	0	0	0	0	0	**1**	**−1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	9	3PGri	0	0	0	0	0	0	0	**1**	0	**1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	10	2PGri	0	0	0	0	0	0	0	0	0	0	**−1**	**−1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	11	PEP	0	0	0	0	0	0	0	0	0	0	0	**1**	**−1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	12	ATP	0	**−1**	0	**−1**	0	0	0	**1**	0	0	0	0	**1**	0	0	**−1**	**1**	0	0	0	0	0	0	0	0	**−1**	0	0	0	0
	13	ADP	0	**1**	0	**1**	0	0	0	**−1**	0	0	0	0	**−1**	0	0	**1**	**−2**	0	0	0	0	0	0	0	0	0	0	0	0	0
	14	6PGlcA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**−1**	0	0	0	0	0	0	0	0	0	0	0
	15	NADP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	**−1**	**−1**	**1**	0	0	0	0	0	0	0	0	0	0
	16	GSH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**2**	**−2**	0	0	0	0	0	0	0	0	0
	17	Rul5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	**−1**	**−1**	0	0	0	0	0	0	0
	18	Xul5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	**−1**	0	0	**−1**	0	0	0
	19	Rib5P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**−1**	0	**−1**	0	0	0	0
	20	Sed7P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**−1**	0	0	0	0	0
	21	E4P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	**−1**	0	0	0
	22	NAD	0	0	0	0	0	0	**1**	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	23	Pi	0	0	0	0	0	0	**1**	0	0	**1**	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0
	24	Lac	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0
	25	Pyr	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**−1**	**−1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**

[a] Non-zero elements are marked in bold.

much lower capacity. One enzyme, the lactate dehydrogenase, is the master enzyme of two different reactions,[60] the oxidation of lactate to pyruvate under formation of $NADH_2$ and $NADPH_2$.

We have determined the possible *k*-subnetworks ($k=1, 2,\ldots,28$) of the complete network shown in Scheme 1 by solving the minimum problem (15) with respect to all enzymes. These computations were performed by using the CPLEX programme (http://www.ilog.com/products/cplex/). The value of the parameter β, which measures the costs of partial fluxes catalysed by assistant enzymes, was arbitrarily put to $\beta=100$. By measuring all fluxes in units of $\mathrm{mM\,h^{-1}}$, the value of the cost factor λ_0 for enzyme usage according to (18) was put to $\lambda_0=30\times100\times3.6\times10^2\ \mathrm{mM\,h^{-1}}=1.08\times10^6\ \mathrm{mM\,h^{-1}}$ to guaranty fulfilment of condition (16). As the actual share of assistant enzymes in the accomplishment of fluxes is difficult to assess and cannot be backed up by experimental data, a second series of computations was performed under omission of the partial fluxes catalysed by assistant enzymes. In this case it was sufficient to choose $\lambda_0=1.08\times10^4\ \mathrm{mM\,h^{-1}}$. The results obtained under inclusion of partial fluxes through assistant enzymes are given in parenthesis in Table 3 and in the following comments. Eleven (ten) enzymes turn out to be essential, that is, no solution of problem (15) can be found if the decision variable for one of these enzymes is put to zero. As mentioned before, enzyme E_9 (bisphosphoglycerate mutase) is not strictly essential as the conversion of 1,3- to 2,3-P_2G can be also catalysed by enzyme E_{11}. Thus, enzyme E_9 becomes essential if the flux through assistant enzymes is excluded. The remaining 18 non-essential enzymes constitute eight (nine) different *k*-subnetworks listed in Table 3 and illustrated in Scheme 2 with ascending values of the flux-evaluation function.

To check the kinetic feasibility of the found subnetworks we used a comprehensive mathematical model[10] describing the kinetics of the erythrocyte network in Scheme 1. Kinetic feasibility of a proposed subnetwork was assessed by putting the v_{max} values of the omitted enzymes (forming the *k*-enzyme cluster of the subnetwork) to zero in the model equations, constraining the fluxes through the four target reactions to the values given above and trying to determine a stationary solution of this system by means of a global nonlinear regression method (Large-scale solver, http://www.frontsys.com). If a locally stable stationary solution was found, the subnetwork was classified as kinetically feasible. This regression calculation was performed by decreasing the maximal activities of the disabled enzymes in a step-wise manner in order to avoid the numerical procedure's starting too far from the attractor of a stable steady state. Although it cannot be guarantied that stable stationary solutions can always be found by this strategy, the inability to reach a stable kinetic steady state for five subnetworks can be made plausible by kinetic arguments (see below). The calculations have shown that stable stationary solutions of the kinetic model exist for only three of the eight subnetworks predicted. Intriguingly, these three kinetically feasible subnetworks exhibit the lowest values of the flux-evaluation function (panels A–C in Scheme 2). Note that the kinetically feasible subnetwork B represents the enzyme-minimised network lacking three enzymes and four reactions.

Table 2. Reactions and enzymes of the erythrocyte network.

v_n	Reaction formula	K^{equ}	Master enzyme	Assistant enzymes	E_n	Enzyme name	Abbr.	EC no.
v_1	Glc(out)→Glc	1.00×10^{0}	E_1		E_1	glucose transporter	GlcT	
v_2	Glc + ATP → Glc6P + ADP	3.90×10^{3}	E_1		E_2	hexokinase	HK	2.7.1.1
v_3	Fru6P → Glc6P	2.55×10^{0}	E_2		E_3	phosphohexose isomerase	GPI	5.3.1.9
v_4	Fru6P + ATP → Fru1,6P_2 + ADP	1.00×10^{5}	E_3		E_4	phosphofructokinase	PFK	2.7.1.11
v_5	DHAP + GraP → Fru1,6P_2	8.77×10^{0}	E_4		E_5	aldolase	ALD	4.1.2.13
v_6	GraP → DHAP	2.46×10^{1}	E_5		E_6	triosephosphate isomerase	TPI	5.3.1.1
v_7	GraP + Pi + NAD → 1,3P_2G + NADH	5.21×10^{3}	E_6		E_7	triosephosphate dehydrogenase (NAD)	GAPDH	1.2.1.12
v_8	1,3P_2G + ADP → 3P_2G + ATP	1.46×10^{3}	E_7		E_8	phosphoglycerate kinase	PGK	2.7.2.3
v_9	1,3P_2G → 2,3P_2G	$\mathbf{1.00\times10^{5}}$	E_9	E_{11}	E_9	bisphosphoglycerate mutase	DPGM	5.4.2.4
v_{10}	2,3P_2G → 3PG + Pi	1.00×10^{5}	E_{10}	E_9	E_{10}	bisphosphoglycerate phosphatase	DPGase	3.1.3.13
v_{11}	2PG → 3PG	6.90×10^{0}	E_{11}	E_9	E_{11}	phosphoglycerate mutase	PGM	5.4.2.1
v_{12}	2PG → PEP	1.70×10^{0}	E_{12}		E_{12}	enolase	EN	4.2.1.11
v_{13}	PEP + ADP → Pyr + ATP	1.38×10^{4}	E_{13}		E_{13}	pyruvate kinase	PK	2.7.1.40
v_{14}	Pyr + NADH → Lac + NAD	9.09×10^{3}	E_{14}		E_{14}	lactate dehydrogenase	LDH	1.1.1.28
v_{15}	Pyr + NADPH → Lac + NADP	1.42×10^{3}	E_{14}		E_{15}	ATPase (total)	ATPase	
v_{16}	ATP → ADP + Pi	$\mathbf{1.00\times10^{5}}$	E_{15}		E_{16}	myokinase (adenylate kinase)	AK	2.7.4.3
v_{17}	2 ADP → ATP + AMP	4.00×10^{0}	E_{16}		E_{17}	glucose-6-phosphate dehydrogenase	G6PD	1.1.1.49
v_{18}	Glc6P + NADP → 6PG + NADPH	2.00×10^{3}	E_{17}		E_{18}	phosphogluconate dehydrogenase	6PGD	1.1.1.44
v_{19}	6PG + NADP → Ru5P + CO_2 + NADPH	1.42×10^{2}	E_{18}		E_{19}	glutathione reductase	GSSGR	1.8.1.7
v_{20}	GSSG + NADPH → 2 GSH + NADP	1.04×10^{0}	E_{19}		E_{20}	glutathione oxidation (total)	GSHox	
v_{21}	2 GSH → GSSG	$\mathbf{1.00\times10^{5}}$	E_{20}		E_{21}	phosphoribulose epimerase	EP	5.1.3.1
v_{22}	Ru5P → X5P	2.70×10^{0}	E_{21}		E_{22}	ribose phosphate isomerase	KI	5.3.1.6
v_{23}	Ru5P → R5P	3.00×10^{0}	E_{22}		E_{23}	transketolase	TK	2.2.1.1
v_{24}	X5P + R5P → GraP + S7P	1.05×10^{0}	E_{23}		E_{24}	transaldolase	TA	2.2.1.2
v_{25}	S7P + GraP → E4P + Fru6P	1.05×10^{0}	E_{24}		E_{25}	phosphoribosylpyrophosphate synthetase	PRPPS	2.7.6.1
v_{26}	R5P + ATP → AMP + PrPP	$\mathbf{1.00\times10^{5}}$	E_{25}		E_{26}	phosphate transporter	PT	
v_{27}	X5P + E4P → GraP + Fru6P	1.20×10^{0}	E_{23}		E_{27}	lactate exchange	LacT	
v_{28}	Pi(out)→Pi	1.00×10^{0}	E_{25}		E_{28}	pyruvate exchange	PyrT	
v_{29}	Lac(out)→Lac	1.00×10^{0}	E_{26}					
v_{30}	Pyr(out)→Pyr	1.00×10^{0}	E_{27}					
v_{31}	PrPP(out)→PrPP	1.00×10^{0}	E_{28}					

Subnetwork A gives the lowest value of the flux-evaluation function (= flux-minimised network). It is characterised by a lack of pyruvate-exchange flux ($d_{28}=0$, $v_{30}=0$). As a consequence, the flux through the $NADPH_2$-dependent lactate dehydrogenation must vanish ($v_{15}=0$). Setting the v_{max} values of the pyruvate exchange and of the $NADPH_2$-dependent activity of the lactate dehydrogenase to zero, the kinetic model provides a stable stationary solution. It has to be noted, however, that putting the activity of the $NADPH_2$-dependent activity of the lactate dehydrogenase to zero appears to be in conflict with the fact that the flux through the $NADH_2$-dependent activity of the lactate dehydrogenase has to be equal to the flux through the glyceraldehydes-3-phosphate dehydrogenase to balance the formation and utilisation of $NADH_2$. Although the two fluxes catalysed by one and the same enzyme (lactate dehydrogenase) can be differently regulated, it should not be possible to drop one flux down to zero whilst the other has to be kept at a finite value. The problem is that our theoretical approach does not include side constraints that force the flux to be different from zero if the catalysing enzyme is present. A reasonable mathematical formulation of this obvious implication is difficult because it requires an a priori assumption about the size of the non-zero fluxes to be made.

In subnetwork B, the flux through the reactions of the non-oxidative pentose pathway is suppressed because the enzymes E_{21} (phosphoribulose epimerase), E_{23} (transketolase) and E_{24} (transaldolase) are disabled. The kinetic simulations show that such a reduced network may indeed operate under normal physiological conditions, in that the flux through the oxidative pentose pathway equals the flux of PRPP synthesis. For the in vivo values of the load parameters chosen in our calculations, the flux of PRPP synthesis ($v_{26}=0.026$ mmol h^{-1}) is lower than the oxidative load ($v_{21}=0.093$ mmol h^{-1}). The flux through the oxidative pentose pathway ($v_{18}=v_{19}$) must equal the flux of PRPP synthesis (v_{26}) and thus is not sufficient to provide all $NADPH_2$ utilised by the glutathione reductase reaction. This residual amount of $NADPH_2$ is produced by the $NADPH_2$-dependent activity of the lactate dehydrogenase working in the backward direction. In this subnetwork, the LDH works effectively as a transhydrogenase: $NADH_2 + NADP \rightarrow NADPH_2 + NAD$.

Subnetwork C has no oxidative pentose pathway, $NADPH_2$ is exclusively delivered by the $NADPH_2$-dependent activity of the lactate dehydrogenase. The kinetic model has no stationary solution in this case. This is plausible considering that the maximal activity of the $NADPH_2$-dependent LDH is $v_{max_{LDH}}=243$ mmol h^{-1} and the equilibrium constant amounts to $K^{equ}_{LDH}=14\,182$ so that, according to the general rate law (10), the magnitude of the back flux yielding $NADPH_2$ cannot be larger than $v_{max}/K^{equ}=0.017$ mmol h^{-1}. Reducing the oxidative load to 10% of the in vivo value, that is, setting $v_{21}=0.0093$ mmol h^{-1}, the

Table 3. *Calculated reduced erythrocyte networks.*

Generating enzyme	Variant	k-Enzyme cluster	k-Reaction cluster	Θ_v
1	H	E_1, E_2, (E_{10}, E_{11})	v_1, v_2	93.47 (382.63)
2	H	E_1, E_2, (E_{10}, E_{11})	v_1, v_2	93.47 (382.63)
3	E	E_3, (E_{10}, E_{11})	v_3	17.06 (306.22)
4	G	E_4, E_5, E_6, (E_{10}, E_{11})	v_4, v_5, v_6	44.95 (334.11)
5	G	E_4, E_5, E_6, (E_{10}, E_{11})	v_4, v_5, v_6	44.95 (334.11)
6	G	E_4, E_5, E_6, (E_{10}, E_{11})	v_4, v_5, v_6	44.95 (334.11)
7	essential			
8	essential			
9	essential (I)	(E_4, E_5, E_6, E_9)	(v_4, v_5, v_6)	(301.26)
10	essential (G)	(E_4, E_5, E_6, E_{10}, E_{11})	(v_4, v_5, v_6)	(334.11)
11	essential (G)	(E_4, E_5, E_6, E_{10}, E_{11})	(v_4, v_5, v_6)	(334.11)
12	essential			
13	essential			
14	essential			
15	essential			
16	essential			
17	C	E_{17}, E_{18}, (E_{12}, E_{13})	v_{18}, v_{19}	12.21 (301.37)
18	C	E_{17}, E_{18}, (E_{12}, E_{13})	v_{18}, v_{19}	12.21 (301.37)
19	essential			
20	essential			
21	B	(E_{10}, E_{11}), E_{21}, E_{23}, E_{24}	v_{22}, v_{24}, v_{25}, v_{27}	12.10 (301.26)
22	D	(E_{10}, E_{11}), E_{22}	v_{23}	12.53 (301.69)
23	B	(E_{10}, E_{11}), E_{21}, E_{23}, E_{24}	v_{22}, v_{24}, v_{25}, v_{27}	12.10 (301.26)
24	B	(E_{10}, E_{11}), E_{21}, E_{23}, E_{24}	v_{22}, v_{24}, v_{25}, v_{27}	12.10 (301.26)
25	essential			
26	essential			
27	F	(E_{10}, E_{11}), E_{27}	v_{29}	21.72 (310.88)
28	A	(E_{10}, E_{11}), E_{28}	v_{15}, v_{30}	12.05 (301.22)

[a] In brackets: disabled enzymes, disabled reactions and flux-evaluation values if catalysis by assistant enzymes is allowed.

kinetic model indeed provided a stationary solution. Thus, the proposed subnetwork C is only conditionally feasible from the kinetic view point, since cancellation of the oxidative pentose pathway can be compensated for by the $NADPH_2$-dependent activity of the lactate dehydrogenase only in the case of a very low oxidative load.

Subnetwork D has no ribose phosphate isomerase (E_{22}), so ribulose-5-phosphate cannot be converted into ribose-5-phosphate. As a consequence, the synthesis of ribose phosphate for the synthesis of PRPP must proceed exclusively along the nonoxidative pentose pathway. For balance reasons, a surplus of ribose-5-phosphate remains that can only be compensated for by directing the flux through the oxidative pentose pathway in the backward direction. The kinetic calculations show that the magnitude of the required backward flux exceeds the maximal backward capacity of the oxidative pentose pathway.

In subnetwork E, the glycolytic flux is interrupted because of lacking phosphohexose isomerase (E_3). In principle, glycolytic traffic can be bypassed through the pentose cycle constituted by the oxidative and nonoxidative pentose pathway. The kinetic calculations show, however, that the flux through the oxidative pentose pathway cannot attain the required magnitude because this flux is strictly controlled by the $NADP/NADPH_2$ ratio. Thus, without the presence of a severe oxidative load stimulating the flux through the oxidative pentose pathway at least 30-fold, this subnetwork cannot work from the kinetic point of view.

Subnetwork F has no lactate exchange. Thus, all lactate formed by glycolysis has to be converted into pyruvate by the $NADPH_2$-dependent lactate dehydrogenase. This is a kinetically infeasible situation because the resulting high flux of $NADPH_2$ production exceeds the maximal backward capacity of the oxidative pentose pathway.

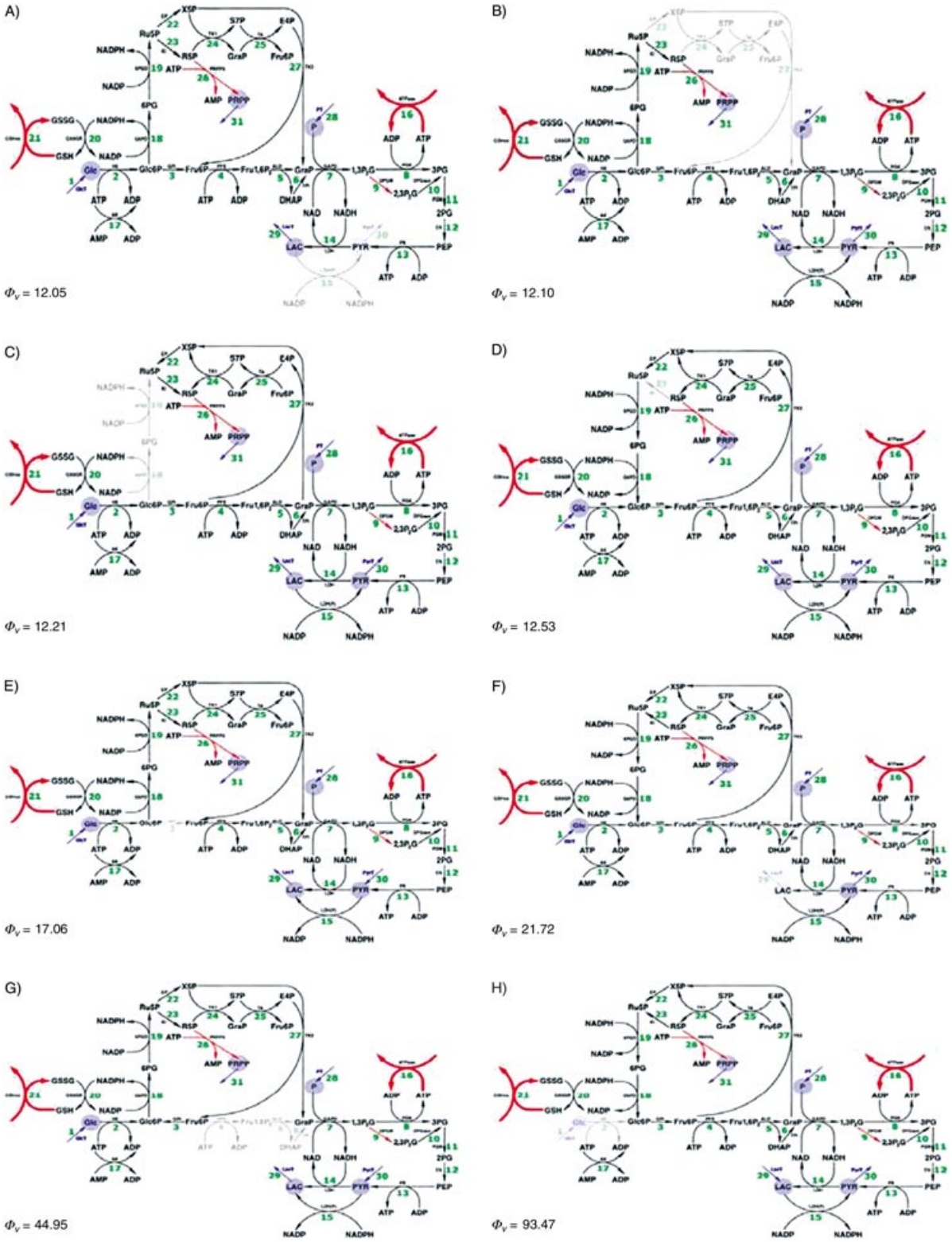
In subnetwork G, the middle part of the glycolytic pathway cannot operate because the enzymes E_4–E_6 are disabled. Analogously to in subnetwork E, the glycolytic flux cannot be bypassed through the pentose cycle because of the restriction of the flux through the oxidative pentose pathway by the oxidative load.

Subnetwork H represents a network variant in which the erythrocyte metabolism is fully supplied by the uptake of lactate. In this case, all the energy driving the endergonic reactions is brought about by a massive production of $NADPH_2$ in reaction R_{15}. The required back flux through the oxidative pentose pathway is incompatible with the maximal backward activities of the G6PD and 6PGD. A second kinetic reason leading to the rejection of this subnetwork is that, at normal lactate levels in the blood (about 1–10 mM), the flux of lactate uptake into the cell is much too low.

In summary, only three of the eight subnetworks predicted by our optimisation approach proved to be kinetically feasible. These kinetic subnetworks are characterised by a marginal increase of less than 5% in the flux evaluation term compared with the flux-minimised reference state. Remarkably, the smallest subnetwork, that is, the subnetwork having the lowest number of enzymes and non-zero fluxes, proved to be kinetically feasible (see Scheme 2B). Rejection of the five remaining variants of subnetworks by kinetic arguments was mostly due to the occurrence of additional backward fluxes that according to general rate law (10) are catalysed with ($1/K^{equ}$)-fold lower effective maximal capacity than the corresponding forward fluxes. This finding is in line with the generally accepted opinion that it is almost impossible to accomplish larger backward fluxes through strongly exergonic reactions.

2. Application to a metabolic scheme of the C metabolism in *Methylobacterium extorquens* AMD

As a second, more complex metabolic network, we have studied the central metabolism of *Methylobacterium extorquens* AMD. This bacterium is capable of growth using C1 compounds such as methanol as the only carbon and energy source. The underlying metabolic scheme (Scheme 3A) is similar to that in ref. [22] with slight modifications. In brief, formaldehyde is produced from methanol by the methanol dehydrogenase complex. The formaldehyde may react with two pools of folate compounds, tetrahydrofolate (H_4F) and tetrahydromethanopterin (H_4MPT). Each of the methylene adducts is involved in further reactions. The reaction scheme comprises the following subsystems: formaldehyde metabolism, glycolysis and gluconeogenesis, TCA cycle, pentose phosphate shunt, serine cycle, poly-β-hydroxy buterate (PHB) synthesis, respira-



Scheme 2. *Reaction scheme for predicted reduced subnetworks of the erythrocyte network shown in Scheme 1. Reaction arrows in light grey indicate zero-fluxes due to enzymes excluded (all parts that cannot be read easily represent disabled reactions and are not vital to the scheme). Note that the reaction arrows point into the direction of the calculated steady-state fluxes. A) Kinetic feasibility: yes. Omission of pyruvate exchange implies balanced production and utilization of $NADH_2$ so that flux through LDH(P) vanishes. B) Kinetic feasibility: yes. Omission of the non-oxidative pentose pathway can be compensated for by a higher flux through the oxidative pentose pathway. C) Kinetic feasibility: conditionally yes. At (very) low oxidative load omission of the oxidative pentose pathway can be compensated for by $NADPH_2$-production by the LDH(P). D) Kinetic feasibility: no. Required backward flux through the oxidative pentose pathway (= flux of PRPP synthesis) exceeds maximal capacity of the 6PGD. E) Kinetic feasibility: no. Maximal flux through the oxidative pentose pathway is restricted by the NADP/$NADPH_2$ ratio. F) Kinetic feasibility: no. High production of $NADPH_2$ induced by full lactate consumption of the LDH(P) exceeds maximal backward capacity of 6PGD. G) Kinetic feasibility: no. Maximal flux through the oxidative pentose pathway is restricted by the NADP/$NADPH_2$ ratio. H) kinetic feasibility: no. Required backward flux through the oxidative pentose pathway (≈ rate of ATP synthesis) exceeds maximal capacity of the 6PGD.*

tion and oxidative phosphorylation. The following metabolites can be exchanged with the external medium by free or facilitated diffusion: methanol, CO_2, formate, glycine, serine, succinate, inorganic phosphate and formaldehyde. All reactions and enzymes of the network are given in Table 4. As in the first example, the reactions are notated such that they proceed from the left to the right under standard conditions; this implies that all equilibrium constants are larger than or equal to unity. If available, the values of the equilibrium constants were taken from ref. [21], otherwise they were fixed to the standard values 1 ($\Delta G_j^{(0)} = 0$) and 10^6 ($\Delta G_j^{(0)} = 28.6$ kJ mol^{-1}) for reactions known to proceed either close to or very far from equilibrium, respectively. The stoichiometric matrix relating the 77 metabolites to the 78 reactions and the reaction-specific enzyme sets are shown in Table 5.

A number of metabolites participating in the central metabolism of *Methylobacterium extorquens* serve as precursors of the so-called biomass of the bacterium or are formed during biomass synthesis. The biomass consists mainly of proteins, PHB and higher carbohydrates.[22] The fluxes describing the incorporation of precursor metabolites into the biomass are considered as the target fluxes of the system. They are indicated by red arrows in the reaction scheme. As the stoichiometric proportions with which the precursor metabolites are consumed or produced during biomass production have been determined experimentally,[22] all fluxes connecting the precursor metabolites with the biomass can be expressed through a single flux, the flux of biomass production (v_{78}), multiplied by the corresponding stoichiometric coefficient (see reaction 78 in Table 4).

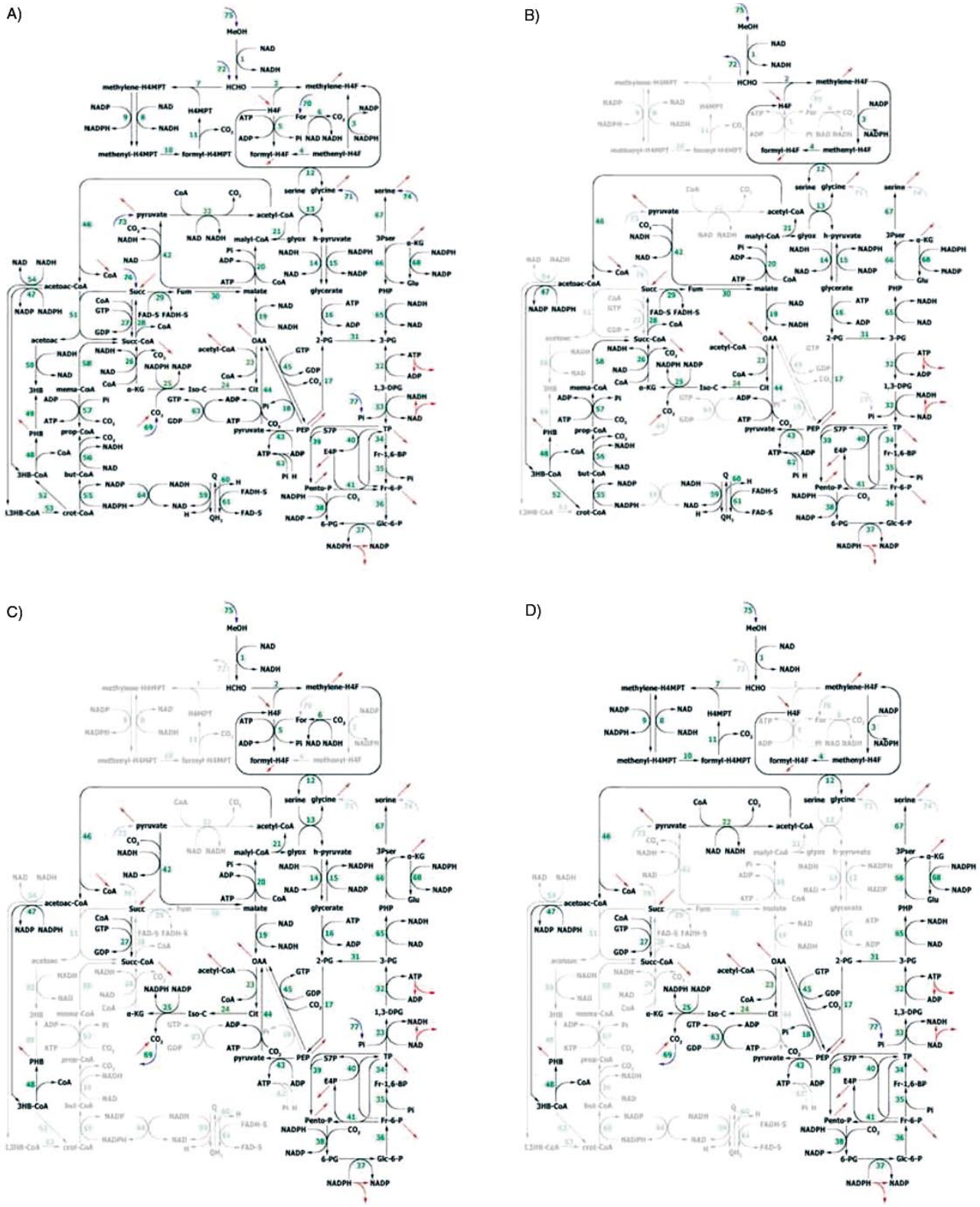
The flux-minimised steady state of the central metabolism of *M. extorquens* in which no pressure is exerted on the economisation of enzymes ($\lambda = 0$) has already been calculated in ref. [17], here methanol was considered to serve as the only available carbon source, that is, the uptake fluxes v_{69}–v_{76} of exchangeable carbon compounds except flux v_{75} (exchange of methanol) were constrained to zero. Intriguingly, 22 fluxes become zero in the flux-minimised state, that is, they are predicted to be dispensable if biomass production is the only function to be accomplished by the central metabolism of the bacterium. Comparison of the calculated fluxes with experimental data from ref. [23], available for 16 (out of 78) reactions (see Figure 1), demonstrates the good overall quality of the predicted flux distribution ($r^2 = 0.68$). For comparison, Figure 1 B and C depict analogous correlation analyses for theoretical flux patterns predicted by two alternative methods of computational flux-pattern analysis that are currently in use. Van Dien et al.[22] were the first to compute the hypothetical flux distribution within a metabolic network of *M. extorquens* that is very similar to that shown in Scheme 3 A. Their approach was based on the calculation of elementary modes and subsequent subjective assignment of flux values to the individual reactions. As shown in Figure 1 B these theoretical flux estimates also agree well with the experimental values, although the correlation coefficient ($r^2 = 0.42$) is smaller than that obtained with the flux-minimisation method. For the sake of completeness, we have also calculated the flux pattern for the metabolic scheme in Scheme 3 A by employing the criterion of maximal biomass production used in several applications of FBA to microorganisms. To this end, we determined the maximum of the flux through reaction R_{78} (biomass production) by constraining the input of methanol to a fixed value and treating all reactions with $K^{equ} < 10^4$. Surprisingly, the resulting flux pattern is not in reasonable accordance with the experimental data (see Figure 1 C), although biomass production is assumed to be the primary objective of this bacterium.

The calculation of subnetworks was performed for two experimental regimes where either methanol or succinate served as the only carbon source (referred to as M and S in the following). For both types of substrates, the minimisation problem (15) was solved for all 77 enzymes. These calculations were run under inclusion of assistant enzymes putting $\beta = 100$ and the cost factor λ_0 for enzyme usage according to (18) to $\lambda_0 = 78 \times 100 \times 3.6 \times 10^2 = 2.81 \times 10^6$ mM h^{-1}. The flux of biomass production was put to the low value of $v_{78} = 1$ mM h^{-1}. After optimisation, all fluxes were rescaled relative to a basis of 10 mol of C1 units entering the system through reaction 1.

The calculations provided 13 (M) and 14 (S) essential enzymes, the subnetworks generated by the nonessential enzymes split into 23 (M) and 22 (S) different variants detailed in Table 6. Each row of the two matrices in Table 6 symbolises the expression vector of a subnetwork. Components marked in light blue indicate dispensable enzymes, yellow components indicate essential enzymes, dark blue components refer to enzymes that proved to be dispensable in all subnetworks. The two last columns of the matrices give the number of disabled enzymes and the value of the flux-evaluation function of the subnetwork.

With methanol as substrate, there are 13 essential enzymes, two of them (E_1 and E_{74}) being trivially essential as they catalyse the import of methanol and the first reaction step in methanol conversion. All enzymes of the so-called serine synthesis (E_{64}–E_{67}) proved to be essential in all subnetworks irrespective of whether methanol or succinate was used as carbon source. This finding is not surprising and can be directly derived from the stoichiometry of the network because the reactions catalysed by these enzymes form a linear segment of the serine synthesising pathway and serine is one essential biomass precursor. Six other enzymes were also predicted to be essential in all subnetworks for both the methanol and the succinate regime. Five of them belong to the TCA cycle and the PHB synthesis, which, next to serine biosynthesis, turn out to be the most essential part of the network regarding the relative number of indispensable enzymes.

A general feature of the calculated subnetworks is that the majority of the disabled enzymes are arranged in clusters that can be assigned to historically defined pathways. Such a concerted disabling of enzymes makes sense as the inactivation of a single enzyme (i.e. putting its decision variable to zero in the optimisation problem (15)) normally makes it impossible to establish the flux-balance condition in the affected pathway. The more linear the affected pathway, that is, the less branching points before or behind the disabled reaction, the more reactions cannot be balanced and thus will sacrifice their enzymes



Scheme 3. *Reaction scheme of the central metabolism of* Methylobacterium extorquens. *Red arrows indicate utilisation or generation of the corresponding metabolite during biomass production, blue arrows indicate exchange fluxes with the external environment, reaction arrows in light grey indicate zero-fluxes due to enzymes excluded (all parts that cannot be read easily represent disabled reactions and are not vital to the scheme). The scheme is based on information outlined in ref. [15] and derived from the KEGG data base (http://www.genome.ad.jp/kegg/). Reactions, enzymes and equilibrium constants are given in Table 4. A) Complete reaction scheme (all fluxes on) of the central metabolism of* M. extorquens. *The reaction arrows point in the direction of the net reaction under standard conditions rather than in the direction taken under standard conditions. B)–D) Reactions graphs for the three groups of subnetworks of* M. extorquens *represented in the upper, middle and lower portion of the matrix in Table 6 and belonging to B) group I, C) group II or D) group III, respectively, when methanol is used a carbon source. The reaction arrows point into the direction of the calculated steady-state fluxes.*

Table 4. *Reactions and enzymes of the metabolic scheme of* Methylobacterium extorquens *AMD.*

	No.	reaction	K_{eq}	master enzyme	assistant enzymes	E_n	enzyme name	EC no
formaldehyde metabolism	1	MeOH + NAD → HCHO + NADH	100 000	E_1		1	methanol dehydrogenase	1.1.1.244
	2	HCHO + H4F → methylene-H4F	100 000	E_2		2	not catalysed	
	3	methenyl-H4F + NADPH → methylene-H4F + NADP	7	E_3		3	methylene H4F dehydrogenase (MtdA)	1.5.1.5
	4	methenyl-H4F → formyl-H4F	2	E_4	E_5	4	methenyl H4F cyclohydrolase	3.5.4.9
	5	formate + ATP + H4F → formyl-H4F + ADP + Pi	41	E_5	E_4	5	formyl H4F synthetase	6.3.4.3
	6	formate + NAD → NADH + CO_2	420	E_6		6	formate dehydrogenase	1.2.1.2
	7	HCHO + H4MPT → methylene-H4 MPT	100 000	E_7		7	formaldehyde-activating enzyme	
	8	methylene-H4MPT + NAD → methenyl-H4MPT + NADH	1	E_8		8	methylene H4MPT dehydrogenase (MtdB)	
	9	methylene-H4MPT + NADP → methenyl-H4MPT + NADPH	1	E_9		9	methylene H4MPT dehydrogenase (MtdA)	n/a
	10	methenyl-H4MPT → formyl-H4MPT	1	E_{10}		10	methenyl H4MPT cyclohydrolase	3.5.4.27
	11	formyl-H4MPT → CO_2 + H4MPT	100 000	E_{11}	E_{12}	11	formyl MFR:H4MPT formyltransferase	1.2.99.5
serine cycle	12	serine + H4F → methylene-H4F + glycine	10	E_{13}		12	formyl MFR dehydrogenase	1.2.99.5
	13	serine + glyox → h-pyruvate + glycine	1	E_{14}		13	serine hydroxymethyltransferase	2.1.2.1
	14	h-pyruvate + NADH → glycerate + NAD	1	E_{15}		14	serine-glyoxylate aminotransferase	2.6.1.45
	15	h-pyruvate + NADPH → glycerate + NADP	1	E_{15}		15	hydroxypyruvate reductase	1.1.1.81
	16	glycerate + ATP → 2-PG + ADP	100 000	E_{16}		16	glycerate kinase	2.7.1.31
	17	PEP → 2-PG	3	E_{17}		17	enolase	4.2.1.11
	18	PEP + CO2 → OAA + Pi	1	E_{18}		18	PEP carboxylase	4.1.1.31
	19	OAA + NADH → malate + NAD	6260	E_{19}		19	malate dehydrogenase	1.1.1.37
	20	malate + CoA + ATP → malyl-CoA + ADP + Pi	100 000	E_{20}		20	malate thiokinase	6.2.1.9
	21	glyox + acetyl-CoA → malyl-CoA	345	E_{21}		21	malyl-CoA lyase	4.1.3.24
TCA cycle	22	pyruvate + NAD + CoA → acetyl-CoA + CO_2 + NADH	100 000	E_{22}		22	pyruvate dehydrogenase	1.2.4.1
	23	acetyl-CoA + OAA → Cit + CoA	100 000	E_{23}		23	citrate synthase	2.3.3.1
	24	iso-C → Cit	14	E_{24}		24	aconitase	4.2.1.3
	25	a-KG + CO_2 + NADPH → iso-C + NADP	1	E_{25}		25	isocitrate dehydrogenase	1.1.1.42
	26	a-KG + NAD + CoA → Succ-CoA + NADH + CO_2	100 000	E_{26}		26	a-KG dehydrogenase	1.2.1.52
	27	Succ + GTP + CoA → Succ-CoA + GDP	2	E_{27}		27	succinyl-CoA synthetase	6.2.1.4
	28	Succ-CoA → Succ + CoA	100 000	E_{28}		28	succinyl-CoA hydrolase	3.1.2.3
	29	Succ + FAD-S → Fum + FADH-S	1	E_{29}		29	succinate dehydrogenase	1.3.5.1
	30	Fum → malate	5	E_{30}		30	fumarase	4.2.1.2
gluconeogenesis and pentose phosphate pathway	31	2-PG → 3-PG	7	E_{31}		31	phosphoglycerate mutase	5.4.2.1
	32	1.3-DPG + ADP → 3-PG + ATP	3226	E_{32}		32	phosphoglycerate kinase	2.7.2.3
	33	1.3-DPG + NADH → TP + NAD + Pi	3	E_{33}		33	glyceraldehyde-3-P dehydrogenase	1.2.1.12
	34	2 TP → Fr-1.6-BP	5555	E_{34}		34	aldolase	4.1.2.13
	35	Fr-1.6-BP → Fr-6P + Pi	174	E_{35}		35	fructose-1,6-bisphosphatase	3.1.3.11
	36	Fr-6P → Glc-6P	2	E_{36}		36	phosphoglucose isomerase	5.3.1.9
	37	Glc-6P + NADP → 6-PG + NADPH	2	E_{37}		37	glucose-6-phosphate dehydrogenase	1.1.1.49
	38	PentoP + CO_2 + NADPH → 6-PG + NADP	14	E_{38}		38	6-phosphogluconate dehydrogenase	1.1.1.44
	39	TP + S7P → 2 PentoP	2	E_{39}		39	transketolase	2.2.1.1
	40	S7P + TP → Ery-4P + Fr-6P	1	E_{40}		40	transaldolase	2.2.1.2
	41	Ery-4P + PentoP → TP + Fr-6P	10	E_{39}		41	malic enzyme	1.1.1.38
	42	malate + NAD → pyruvate + CO_2 + NADH	1	E_{41}		42	pyruvate kinase	2.7.1.40
	43	PEP + ADP → pyruvate + ATP	18 000	E_{42}		43	pyruvate carboxylase	6.4.1.1
	44	pyruvate + CO_2 + ATP → OAA + ADP	7	E_{43}		44	PEP carboxykinase	4.1.1.32
	45	OAA + GTP → PEP + GDP + CO_2	12	E_{44}		45	β-ketothiolase	2.3.1.16

Table 4. (Continued)

	No.	reaction	K_{eq}	master enzyme	assistant enzymes	E_n	enzyme name	EC no
PHB synthesis	46	2 acetyl-CoA → acetoac-CoA + CoA	1	E_{45}		46	acetoacetyl-CoA reductase (NADPH)	1.1.1.36
	47	acetoac-CoA + NADPH → 3 HB-CoA + NADP	1	E_{46}		47	PHB synthase	2.3.1.–
	48	3 HB-CoA → PHB + CoA	100 000	E_{47}		48	PHB depolymerase	3.1.1.75
acetyl CoA conversion pathway	49	PHB → 3 HB	100 000	E_{48}		49	β-hydroxybutyrate dehydrogenase	1.1.1.30
	50	acetoac + NADH → 3 HB + NAD	526	E_{49}		50	acetoacetate-succinyl-CoA transferase	2.8.3.5
	51	acetoac-CoA + Succ → acetoac + Succ-CoA	100	E_{50}		51	D-crotonase	4.2.1.17
	52	crot-CoA → 3 HB-CoA	6	E_{51}		52	L-crotonase	4.2.1.17
	53	L3HB-CoA → crot-CoA	6	E_{52}		53	acetoacetyl-CoA reductase (NADH)	1.1.1.35
	54	acetoac-CoA + NADH → L3HB-CoA + NAD	1587	E_{53}		54	crotonyl-CoA reductase	1.3.1.8
	55	crot-CoA + NADPH → but-CoA + NADP	1	E_{54}		55	unknown pathway	
	56	but-CoA + NAD → prop-CoA + NADH + CO_2	100 000	E_{55}		56	propionyl-CoA carboxylase	6.4.1.3
	57	mema-CoA + ADP + Pi → prop-CoA + CO_2 + ATP	123	E_{56}		57	methylmalonyl-CoA mutase	5.4.99.2
	58	mema-CoA → Succ-CoA	19	E_{57}		58	NADH-quinone oxidoreductase	1.6.99.5
respiration and oxidative phosphorylation	59	NADH + Q → NAD + 2 H + QH_2	1	E_{58}		59	cytochrome oxidase	1.10.2.2
	60	QH_2 → Q + 2 H	1	E_{59}		60	ubiquinone oxidoreductase	1.5.5.1
	61	FADH-S + Q → FAD-S + QH_2	1	E_{60}		61	ATPase	
	62	ADP + Pi + 2 H → ATP	1	E_{61}		62	NDP kinase	2.7.4.6
	63	GDP + ATP → GTP + ADP	1	E_{62}		63	transhydrogenase	1.6.1.2
	64	NADH + NADP → NADPH + NAD	1	E_{63}		64	3-phosphoglycerate dehydrogenase	1.1.1.95
serine synthesis	65	NADH + PHP → 3-PG + NAD	10 000	E_{64}		65	phosphoserine transaminase	2.6.1.52
	66	a-KG + 3Pser → PHP + Glu	7	E_{65}		66	phosphoserine phosphatase	3.1.3.3
	67	3 Pser → serine	673	E_{66}		67	glutamate dehydrogenase	1.4.1.4
	68	a-KG + NADPH → Glu + NADP	10 000 000	E_{67}		68	ex-CO_2	
exchange fluxes	69	CO_2(out) → CO_2	1	E_{68}		69	ex-formate	
	70	formate(out) → formate	1	E_{69}		70	ex-glycine	
	71	glycine(out) → glycine	1	E_{70}		71	ex-HCHO	
	72	HCHO(out) → HCHO	1	E_{71}		72	ex-pyruvate	
	73	pyruvate(out) → pyruvate	1	E_{72}		73	ex-serine	
	74	serine(out) → serine	1	E_{73}		74	ex-MeOH	
	75	MeOH(out) → MeOH	1	E_{74}		75	ex-Succ	
	76	Succ(out) → Succ	1	E_{75}		76	ex-Pi	
	77	Pi(out) → Pi	1	E_{76}		77	biomass production	
biomass	78	2 methylene-H4F + 11 formyl-H4F + 13 glycine + 7 serine + 7-Succ-CoA + 11 PEP + 41 OAA + 53 acetyl-CoA + 42 pyruvate + 17 a-KG + 2 TP + 16 Fr-6P + 10 PentoP + 5 Ery-4P + 93 PHB + 585 ATP +		E_{77}				
production	240	NADPH + 5 NAD → 585 ADP + 240 NADP + 5 NADH + 7 Succ + 60 CoA + 13 H4F + 118 CO_2						

Table 5. *Stoichiometric matrix of the metabolic network of* Methylobacterium extorquens *AMD.*[a]

| metabolites | | fluxes | | |

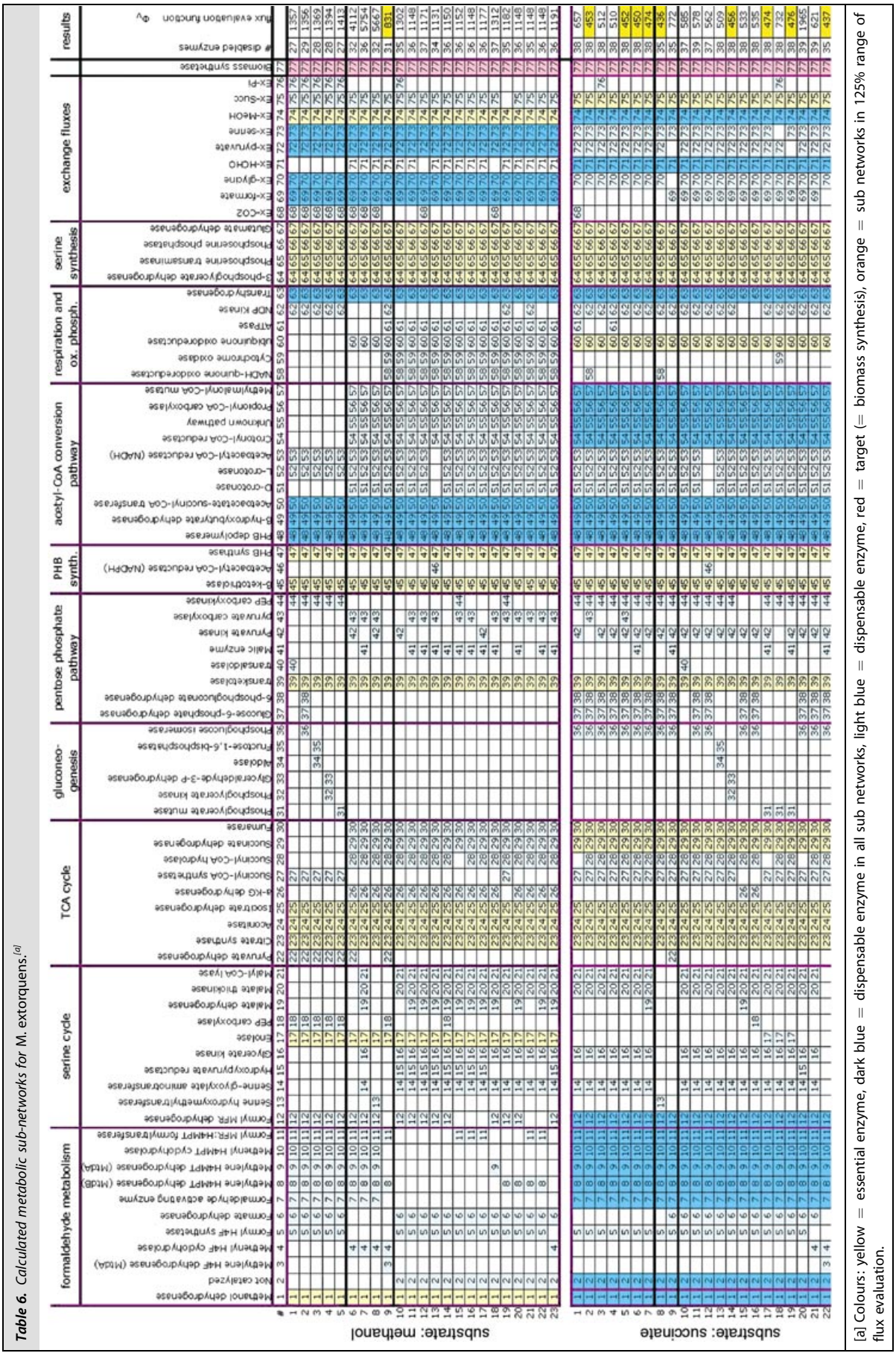


Table 6. *Calculated metabolic sub-networks for* M. extorquens.[a]

[a] Colours: yellow = essential enzyme, dark blue = dispensable enzyme in all sub networks, light blue = dispensable enzyme, red = target (= biomass synthesis), orange = sub networks in 125% range of flux evaluation.

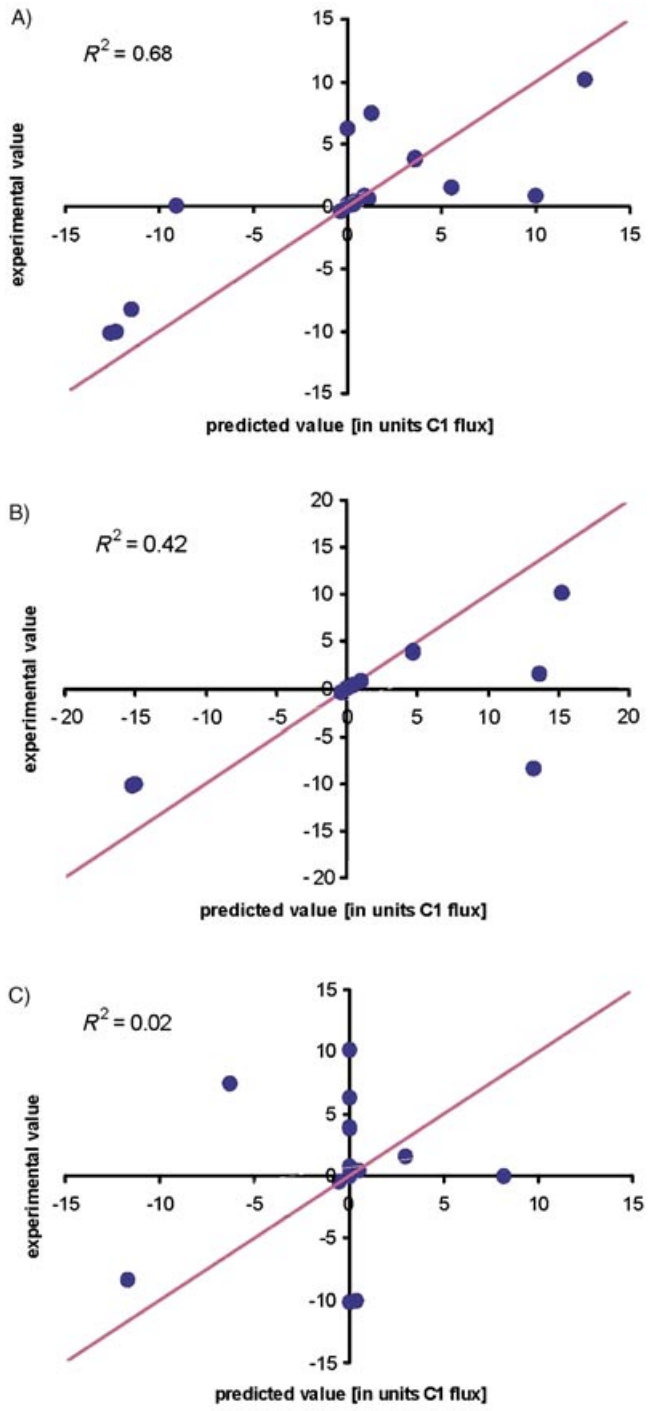


Figure 1. *Correlation between theoretically predicted and measured flux rates for the complete reaction scheme of* M. extorquens. *The experimental values obtained by ^{13}C labelling were taken from ref. [23]. A) Theoretical flux values predicted by flux minimisation (solution of the minimum problem (14) with $\lambda = 0$). B) Theoretical flux values taken from ref. [22]. The calculation of these fluxes was based on an elementary mode analysis of the network and subsequent assignment of flux values according to the frequency with which the individual reactions occurred in the various elementary modes. C) Theoretical flux values calculated by maximising the biomass flux (v_{78}) at fixed exchange flux for methanol and treating all reactions with $K^{equ} > 10^4$ as irreversible. Note that the theoretical flux values have been re-scaled to obtain a 45° linear regression line.*

for economical reasons. Nevertheless, the observed clusters of simultaneously dispensable enzymes cannot be simply inferred from the stoichiometry of the reaction scheme. Creating overall reactions by lumping together all subsequent reactions along a linear (unbranched) segment of the network that trivially carries the same flux at steady-state, the predicted subnetwork 11 (see Table 6), for example, instead of the 41 disabled individual reactions still encompasses 22 disabled overall reactions that form a completely branched subnetwork; this cannot be inferred from the stoichiometric matrix by mere algebraic transformations.

Enzymes predicted to be dispensable in all or most predicted subnetworks are to a larger extent resident in the acetyl-CoA conversion pathway, the formaldehyde metabolism and the membrane exchange fluxes. The conversion of formaldehyde may proceed along two alternative branches containing either methylene–H4MPT or methylene–H4F as intermediate. The branch via methylene–H4F leads to the need for formyl–H4F as a precursor of the biomass production. This may account for the fact that enzymes located in the E_3–E_6 branch of the network are less dispensable than that in the alternative branch (E_7–E_{11}). The acetyl-CoA conversion pathway represents a possible route to synthesise the central metabolite succinyl-CoA. Obviously, the production of succinyl-CoA along the TCA cycle is more favourable from the thermodynamic view point, so enzymes of the acetyl-CoA conversion pathway are dispensable in almost all proposed subnetworks.

Although the subnetworks calculated for the two alternative carbon sources share a lot of similarities, a striking difference between them concerns the enzymes for oxidative energy production. For methanol as substrate, our algorithm predicts that the majority of subnetworks will manage without the respiratory chain and F_0F_1–ATPase. In contrast, for succinate as substrate, the system of aerobic energy production is predicted to be obligatory. These predictions are consistent with the observation that mutants of *M. extorquens* lacking the ubiquinone oxidoreductase showed normal growth on methanol and impaired growth on pyruvate and succinate.[23] Moreover, this finding clearly shows that the apparent dispensability of enzymes also depends upon the concrete environmental conditions of the cell.

Unlike the erythrocyte subnetworks, the subnetworks of *M. extorquens* exhibit a considerable degree of overlap. With methanol as substrate, they can be roughly subdivided into three groups (separated by the horizontal thick black lines in Table 6). Typical representatives of these three groups of subnetworks are depicted in Scheme 3 B–D. Group I comprises subnetworks equipped with a minimum number of enzymes in the formaldehyde metabolism and lacking several enzymes of the acetyl-CoA conversion pathway albeit still capable of synthesising succinyl-CoA from acetoacetyl-CoA along the pathway constituted by the reactions R_{47}, R_{52}, R_{55}, R_{56}, R_{57}, R_{58}. Subnetworks belonging to group II are even more reduced compared with the subnetworks of group I in that they lack some enzymes required for aerobic energy production and completely lack the enzymes of the acetyl-CoA conversion pathway. These subnetworks obtain succinate only as a by-

product of the biomass synthesis. Group III comprises the smallest subnetworks lacking, on top of the enzymes lacking in networks in group II, most enzymes of the serine cycle. The network shown in Scheme 3D represents one of the two enzyme-minimised subnetwork with methanol as substrate. It comprises only 40 enzymes, that is, about 50% of the enzymes should be dispensable at extreme limitations of the protein-synthesising capacity of the cell without compromising the biomass production of *M. extorquens*. The kinetic feasibility of this extreme network could only be assessed by genetic experiments (multiple knock-outs) or on the basis of a validated kinetic model. Neither prerequisite is available at the moment. Based on findings with the subnetworks of erythrocyte metabolism, kinetic feasibility correlates to the value of the flux-evaluation function. For the subnetworks of erythrocyte metabolism, kinetic feasibility tested on the basis of a comprehensive kinetic model was restricted to subnetworks having relative flux-evaluation values of less than 105%. With this threshold value, we had to conclude that none of the subnetworks calculated for *M. extorquens* is kinetically feasible. However, as the relative increase in flux evaluation depends upon the absolute magnitude of the fluxes affected as well as upon the flux distribution as a whole, a slightly higher threshold value of 125% still seems to be reasonable. In that case, one of the 23 calculated subnetworks (9) is expected to be kinetically feasible if methanol is used as substrate. On the other hand, there should be 9 kinetically feasible variants among the 22 subnetworks calculated with succinate as substrate (see Figure 2). Notably, a significant higher number of kinetically feasible subnetworks remains for the succinate regime if the upper threshold value for flux evaluation is varied. This leads to the conclusion that using a multiple carbon source instead of a C1-compound increases the flexibility of the metabolic networks against changes in enzyme activities.

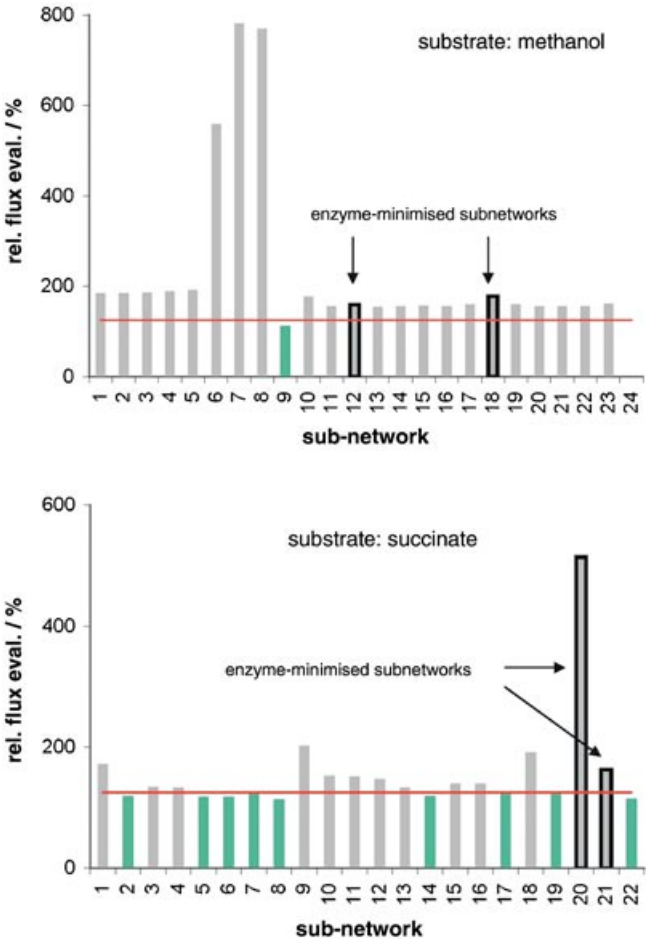


Figure 2. *Relative flux-evaluation value (= % flux-evaluation value of the subnetwork with respect to the flux-evaluation value of the flux-minimised network) for the various subnetworks of* M. extorquens. *The horizontal red line indicates an arbitrarily chosen threshold value of 125% to classify the subnetworks as kinetic feasible (relative flux-evaluation value is smaller than 125%) or unfeasible.*

Discussion

In this work we have applied mathematical optimisation methods to tackle the problem of reducing the number of enzymes in a given metabolic network without compromising its functionality. Our interest in this problem arises from several biological aspects.

First, there is a steadily growing body of publications devoted to the phenomenon of robustness of cellular networks against mutations. This phenomenon is often accounted for by overlapping functions of genes according to which loss-of-function mutations in one gene will have little phenotype effect if there is one or more additional genes with similar functions. However, as correctly pointed out by Wagner[24] and Kitami et al.[25] the more important source of robustness has its origin in the interaction between genes with unrelated functions, or more precisely, in the ability of the cellular reaction network to compensate for the loss of certain reactions by using alternative routes. By putting increasing pressure on reducing the number of enzymes and related reactions, our approach enables such compensation strategies of the cellular metabolism that assure the maintenance of essential output fluxes by various modes of internal flux distributions to be studied.

Secondly, elucidating groups of enzymes that can be conjointly disabled in order to concentrate the protein-synthesising capacity of the cell on more relevant biochemical subsystems offers a way to better rationalise gene-expression profiles monitored by microarray techniques. Typically, genes that appear to be coexpressed or cosuppressed in the microarray analysis cannot be simply attributed to larger functional units of the cellular network. Rather, the expression pattern appears to be rather enigmatic in that certain enzymes of a well-defined pathway are up-regulated and other enzymes of the same pathway are down-regulated. Exactly this type of expression pattern is predicted by our approach and is due to the fact that disabled reactions are bypassed by using parts of or even single reactions of other pathways normally used in a different context.

Thirdly, studying interactions in larger networks on the basis of smaller synthetic networks is a convenient method of hypothesis testing and makes these systems more amenable to

mathematical analysis. For example, unravelling the sequence of whole genomes has strengthened efforts to better understand interaction among genes by creating and analysing networks composed of a limited number of genes.[26,27] Similar analyses based on smaller reconstituted enzyme systems already have a longer tradition in computational studies aimed at elucidating the key modes of regulation in metabolic networks (e.g.,[28,29]). We think that the theoretical approach developed in this paper may serve as a heuristic to design minimal metabolic networks capable of performing a restricted number of specific functions.

Evidently, the ultimate way of studying the possible effects of reduced or completely down-regulated enzyme activities on the stationary states of a metabolic network by computational methods consists of computer simulations based on reliable kinetic models. Unfortunately, mathematical models of this quality are currently available for only a few metabolic pathways with known kinetics of all enzymes involved. Therefore, we have chosen an approach that falls into the field of "structural modelling" going without enzyme-kinetic information. Only the stoichiometry of the system and, if available, some plausible side conditions constraining the external fluxes are used as input. Several methods for the structural (=topological) analysis of metabolic networks have been developed and successfully applied in the last decade. These methods can be grouped into two categories.

One group of methods aims at the identification of basic flux modes of "basic" metabolic operation modes in which only some of the reactions are active. Mathematically speaking, these basic flux modes constitute the edges of the steady-state flux cone defined by the flux-balance conditions and additional (linear) side constraints imposed on the fluxes. Knowledge of these basic flux modes allows metabolic pathways to be redefined in a rigorous quantitative and systemic way[30,31] and the robustness of metabolic networks against insertions or deletion of certain enzymes to be assessed.[32,33] Schuster and co-workers have proposed an algebraic procedure to break down the stationary fluxes in a metabolic network into so-called elementary modes, which they define "as smallest sets of enzymes that can operate at steady state, with all enzymes weighted by the relative flux they need to carry out for the mode to function".[34] These elementary flux modes have strong similarities with the so-called extreme pathways forming a basis in the space of flux distributions restraint by inequality relations.[30] Besides other applications,[35–38] such an extreme pathway analysis has been applied to the metabolism of human red blood cells[39] by using a kinetic scheme similar to that shown in Scheme 1. This analysis has helped to detect functional units in the metabolism of this cell that deviate from the historically defined metabolic pathways. However, a direct comparison of the calculated extreme pathways with the reduced networks predicted in this paper is not possible because extreme flux modes, as well as elementary modes, by definition represent minimal stand-alone parts of the network and thus do comprise neither all target reactions yielding the functionally important products nor all available resource fluxes. Thus physiologically meaningful subnetworks that enable the maintenance of simultaneously required cellular functions represent combinations of extreme flux patterns, and the determination of these linear combinations is beyond the scope of extreme-pathway analysis and requires methods as those developed in this paper.

A second group of methods in the field of structural modelling is commonly referred to as flux-balance analysis (FBA).[13] In contrast to decomposition methods, this method includes the steady-state fluxes of the network and allows all metabolites (including co-enzymes) to be treated in an equal manner. FBA is based on the premise that the most likely distribution of stationary fluxes in the metabolic network has to be optimal with respect to a feasible optimisation criterion. The definition of the optimisation criterion is the key point of the whole approach. It is the common view that the principles governing the design of cells, tissues and organisms can only be grasped in the context of natural evolution. In the Darwinian sense, natural evolution is a permanent optimisation process leading to the survival of phenotypes that are best adapted to their natural environment. With respect to metabolism, best adaptation to environmental conditions may involve multiple properties, like robustness against fluctuations in the supply of external substrates or relative insensitivity to alterations in the structure and function of the underlying proteins (enzymes, transporters). Accordingly, various optimisation criteria have been suggested, including maximisation of steady-state reaction rates[40,41] or the stability of steady states and system responses to parameter changes.[42–45] It is very unlikely that a single evolutionary principle accounts for the sophisticated regulation of metabolic systems of currently existing cells. Thus, resting the computational prediction of system properties on a single optimisation criterion a priori will necessarily hold a considerable degree of arbitrariness.

In previous applications of FBA, the maximisation of biomass production was used as flux-optimisation criterion. Despite numerous applications of FBA[46–56] based on this criterion, a direct demonstration of its feasibility by comparison with measured flux rates was only provided in ref. [57]. On the other hand, our calculations have shown that its application to another relatively simple cell as bacterium *M. extorquens* led to the prediction of a flux pattern that clearly differs from the observed one. It has to be emphasised that this finding by no means compromises the concept of FBA, which traditionally focused on microorganisms that had growth as their main objective. It simply indicates the need to look for alternative optimisation criteria, in particular for studying the metabolism of multifunctional vertebrate cells, such as hepatocytes or nerve cells, where the maximal production of biomass can hardly be taken as an appropriate optimisation criterion.

To overcome the obvious shortcomings associated with using biomass production as the only ultimate goal of cellular metabolism, the minimisation of total internal fluxes was proposed as an alternative to account for the flux distribution in metabolic networks at steady state.[17] This principle captures the obvious fact that gaining functional fitness with minimal expense of external resources and along the shortest route in the evolutionary landscape must have been a decisive selec-

tion factor during the natural evolution of cellular systems. In this paper, we have adopted this concept and extended it by defining an objective function that measures two different kinds of expenses that arise when keeping a metabolic network in operation. On one hand, expenses arise for the energy and external metabolites needed to keep the participating enzymes on a sufficiently high level. In our approach, these expenses are measured by a single parameter, the cost factor λ. For the sake of simplicity, we have ignored the fact that maintaining the cellular concentration of short lived proteins will certainly require more energy for protein synthesis and ATP-dependent proteolysis than maintaining the concentration of proteins with long half-life. Thermodynamic expenses arise from establishing a nonequilibrium state, that is, accomplishing nonvanishing fluxes through the various reactions of the network. These expenses are measured in terms of a flux-evaluation function representing the weighted sum of all fluxes.

Weighting of the fluxes is necessary for two reasons. First, establishing a sufficiently high flux through a given reaction by the side activity of an assistant enzyme is more difficult than by using a master enzyme. In our approach this fact was taken into account by putting a higher weight ($\beta > 1$) on fluxes catalysed by an assistant enzyme. It has to be mentioned that a systematic study of the reaction, specificity of known enzymes, has been undertaken only in a few cases, so their side activities are often not known. In the examples considered in this paper, assistant enzymes played a minor role. It is conceivable however, that increasing experimental knowledge about the whole spectrum of catalytic capabilities of enzymes will make this part of our theoretical approach an essential prerequisite for fully understanding the striking flexibility of cellular metabolism in the face of external perturbations. Secondly, reversing the direction of fluxes with respect to the "natural" forward direction defined at standard conditions becomes more and more unfavourable from the thermodynamic viewpoint, the larger the thermodynamic equilibrium constant of the reaction. Differential weighting of the forward and backward fluxes takes into account that the increase in the concentration of the reaction products necessary to reverse the "natural" direction of a reaction is proportional to the thermodynamic equilibrium constant K^{equ}, which, for larger values of K^{equ}, should give rise to an unfavourable osmotic balance. As suggested by Atkinson,[58] living cells encompass myriads of metabolites, while their solvent capacity is limited and the internal osmotic pressure must not exceed certain critical values. Is has to be emphasised that the standard free-energy change ΔG_0 or the thermodynamic equilibrium constant K^{equ} may serve only as a crude measure of the true thermodynamics of a reaction because of the strong dependence of the free-energy change from the (unknown) concentrations of the reactants [Eq. (8)]. For example, the standard free-energy change for ATP hydrolysis ($ATP + H_2O \rightarrow ADP + P_i$) amounts to about $-30\ kJ\,mol^{-1}$ ($K^{equ} = 1.6 \times 10^5$), whereas under cellular conditions with $[ATP] \approx 1$ mM, $[P_i] \approx 1$ mM and $[ADP] \approx 0.1$ mM, the actual value of the free-energy change is about $-50\ kJ\,mol^{-1}$, corresponding to an apparent equilibrium constant of about 4.9×10^8, that is about three orders of magnitude higher than K^{equ}.

Defining the objective function as the sum of expenses for enzyme synthesis and flux accomplishment, we postulate the most reliable metabolic network to be given by the minimum of the objective function at the side constraints that the flux balance relations are met with respect to all internal metabolites and that those target fluxes directly linked to the physiological functions of the cell have to assume prescribed values. In order to generate various variants of enzyme-reduced subnetworks, we have determined the minimum of the objective function under the additional constraint that the activity of an arbitrary enzyme—the so-called generating enzyme—is put to zero. Constraining the activity of a single enzyme to zero may even imply that no stationary flux distribution can be established anymore. In this case, the generating enzyme is said to be essential.

The proposed mathematical approach was applied to identify reduced metabolic subnetworks in two metabolic networks of different complexity, the energy and redox metabolism of erythrocytes and the central carbon metabolism of the *M. extorquens*. For the metabolic scheme of the erythrocyte, a comprehensive and validated kinetic model was available that allowing the kinetic feasibility of the predicted subnetworks to be assessed, that is, to check whether a locally stable stationary solution of the kinetic model can be obtained if the maximal activity of all enzymes predicted to be dispensable is put to zero. This analysis revealed that the likelihood of a subnetwork's being kinetically feasible decreases with increasing flux-evaluation value. High flux-evaluation values indicate that one or more endergonic reactions would have to proceed in the backward direction to bypass disabled reactions. As the maximal backward flux is limited by v_{max}/K^{equ}, this turns out to be impossible for strongly exergonic reactions. Intriguingly, the flux-evaluation value (and thus the likelihood of failing the criterion of kinetic feasibility) is not strictly proportional to the number of disabled enzymes. For the erythrocyte network, the enzyme-minimised subnetwork (lacking three enzymes and four reactions) is among those subnetworks that proved to be kinetically feasible.

Our analysis provided evidence that 14 out of 28 enzymes (= 50%) of the erythrocyte are essential. In the metabolic network of *M. extorquens* 13 (for methanol as substrate) and 14 (for succinate as substrate) enzymes out of 77 ($\approx$17%) are predicted to be essential. Obviously, the relative portion of essential enzymes decreases with increasing connectivity of the network if network connectivity is quantified by the average number of reactions associated with a single metabolite, that is, the average number of non-zero elements in the rows of the stoichiometric matrix. This connectivity measure amounts to 2.94 and 4.18 for the erythrocyte and bacterial networks, respectively. A fraction of about 17% essential enzymes theoretically predicted for the network of *M. extorquens* is in good agreement with experimental results obtained by systematic genetic analysis with ordered arrays of yeast deletion mutants, according to which more than 80% of the approximately $\approx$6200 predicted genes in *Saccharomyces cerevisiae* are nonessential.[33] It has to be noted that our approach yields a lower boundary for the number of essential enzymes because kinetic

constraints not considered in our approach may only further increase this number.

Inspection of the subnetworks predicted for the metabolic network of *M. extorquens* leads to the following main conclusions. i) The type and number of dispensable enzymes depends upon the external conditions of the cell. A similar conclusion was drawn by Papp et al.,[59] who analysed the dispensability of enzymes in yeast using an in silico flux model. For example, we found striking differences in the usage of enzymes of aerobic energy production with either methanol or succinate as metabolic input. ii) Disabled enzymes form groups that can be related to well-defined metabolic pathways. However, only in a few cases are all enzymes of a "textbook pathway" simultaneously disabled. In most cases, only a fraction of enzymes is dispensable whereas others are still required to allow a flux through a certain part of the pathway. Such patterns resemble those typically seen in microarray gene-expression studies.

Enzymes

Hexokinase [2.7.1.1]; phosphohexose isomerase [5.3.1.9]; phosphofructokinase [2.7.1.11]; aldolase [4.1.2.13]; triosephosphate isomerase [5.3.1.1]; glyceraldehyde-3-phosphatedehydrogenase [1.2.1.12]; phosphoglycerate kinase [2.7.2.3];bisphosphoglycerate mutase [5.4.2.4]; bisphosphoglycerate phosphatase [3.1.3.13]; phosphoglycerate mutase [5.4.2.1]; enolase [4.2.1.11]; pyruvate kinase [2.7.1.40]; lactate dehydrogenase [1.1.1.28]; adenylate kinase [2.7.4.3]; glucose 6-phosphate dehydrogenase [1.1.1.49]; phosphogluconate dehydrogenase [1.1.1.44]; glutathione reductase [1.8.1.7]; phosphoribulose epimerase [5.1.3.1]; ribose phosphate isomerase [5.3.1.6]; transketolase [2.2.1.1]; transaldolase [2.2.1.2]; phosphoribosylpyrophosphate synthetase [2.7.6.1]; transketolase [2.2.1.1]; ethanol dehydrogenase [1.1.1.244]; methylene H4F dehydrogenase (MtdA)[1.5.1.5]; ; methenyl H4F cyclohydrolase [3.5.4.9]; formyl H4F synthetase [6.3.4.3]; formate dehydrogenase [1.2.1.2]; formaldehyde activating enzyme [unknown1; methylene H4MPT dehydrogenase (MtdB) [unknown]; methylene H4MPT dehydrogenase (MtdA) [unknown]; methenyl H4MPT cyclohydrolase [3.5.4.27]; formyl MFR:H4MPT formyltransferase [unknown]; formyl MFR dehydrogenase [1.2.99.5]; serine hydroxymethyltransferase [2.1.2.1]; serine-glyoxylate aminotransferase [2.6.1.45]; hydroxypyruvate reductase [1.1.1.81]; glycerate kinase [2.7.1.31]; PEP carboxylase [4.1.1.31]; malate dehydrogenase [1.1.1.37]; malate thiokinase [6.2.1.9]; malyl-CoA lyase [4.1.3.24]; pyruvate dehydrogenase [1.2.4.1]; citrate synthase [2.3.3.1]; aconitase [4.2.1.3]; isocitrate dehydrogenase [1.1.1.42]; α-KG dehydrogenase [1.2.1.52]; succinyl-CoA synthetase [6.2.1.4]; succinyl-CoA hydrolase [3.1.2.3]; succinate dehydrogenase [1.3.5.1]; fumarase [4.2.1.2]; malic enzyme [1.1.1.38]; pyruvate carboxylase [6.4.1.1]; PEP carboxykinase [4.1.1.32]; β-ketothiolase [2.3.1.16]; acetoacetyl-CoA reductase (NADPH) [1.1.1.36]; PHB synthase [2.3.1.–]; PHB depolymerase [3.1.1.75]; β-hydroxybutyrate dehydrogenase :[1.1.1.30]; acetoacetate-succinyl-CoA transferase [2.8.3.5]; D-crotonase [4.2.1.17]; L-crotonase [4.2.1.17]; acetoacetyl-CoA reductase (NADH) [1.1.1.35]; crotonyl-CoA reductase [1.3.1.8]; propionyl-CoA carboxylase [6.4.1.3]; methylmalonyl-CoA mutase [5.4.99.2]; NADH-quinone oxidoreductase [1.6.99.5]; cytochrome oxidase [1.10.2.2]; ubiquinone oxidoreductase [1.5.5.1]; NDP Kinase [2.7.4.6]; transhydrogenase [1.6.1.2]; 3-phosphoglycerate dehydrogenase [1.1.1.95]; phosphoserine transaminase [2.6.1.52]; phosphoserine phosphatase [3.1.3.3]; glutamate dehydrogenase [1.4.1.4].

Keywords: cellular functions · kinetics · metabolic networks · redundancy · systems biology · theory

[1] J. L. DeRisi, V. R. Iyer, P. O. Brown, *Science* **1997**, *278*, 680 – 686.
[2] R. Korke, L. Gatti, A. L. Lau, J. W. Lim, T. K. Seow, M. C. Chung, W. S. Hu, *J. Biotechnol.* **2004**, *107*, 1 – 17.
[3] A. Zaslaver, A. E. Mayo, R. Rosenberg, P. Bashkin, H. Sberro, M. Tsalyuk, M. G. Surette, U. Alon, *Nat. Genet.* **2004**, *36*, 486 – 491.
[4] E. Klipp, R. Heinrich, H. G. Holzhütter, *Eur. J. Biochem.* **2002**, *269*, 5406 – 5413.
[5] T. A. Rapoport, R. Heinrich, *Biosystems* **1975**, *7*, 120 – 129.
[6] M. Schauer, R. Heinrich, S. M. Rapoport, *Acta Biol. Med. Ger.* **1981**, *40*, 1659 – 1682.
[7] M. Schauer, R. Heinrich, S. M. Rapoport, *Acta Biol. Med. Ger.* **1981**, *40*, 1683 – 1697.
[8] R. Heinrich, *Biomed. Biochim. Acta* **1985**, *44*, 913 – 927.
[9] R. Schuster, H. G. Holzhütter, G. Jacobasch, *Biosystems* **1988**, *22*, 19 – 36.
[10] R. Schuster, H. G. Holzhütter, *Eur. J. Biochem.* **1995**, *229*, 403 – 418.
[11] N. Jamshidi, J. S. Edwards, T. Fahland, G. M. Church, B. O. Palsson, *Bioinformatics* **2001**, *17*, 286 – 287.
[12] B. Teusink, J. Passarge, C. A. Reijenga, E. Esgalhado, C. C. van der Weijden, M. Schepper, M. C. Walsh, B. M. Bakker, K. van Dam, H. V. Westerhoff, J. L. Snoep, *Eur. J. Biochem.* **2000**, *267*, 5313 – 5329.
[13] A. Varma, B. O. Palsson, *Bio/Technol* **1994**, *12*, 994 – 998.
[14] A. Cornish-Bowden, M. L. Cardenas, *Nature* **2002**, *420*, 129 – 130.
[15] C. H. Schilling, J. S. Edwards, B. O. Palsson, *Biotechnol. Prog.* **1999**, *15*, 288 – 295.
[16] J. S. Edwards, R. U. Ibarra, B. O. Palsson, *Nat. Biotechnol.* **2001**, *19*, 125 – 130.
[17] H. G. Holzhütter, *Eur. J. Biochem.* **2004**, *271*, 2905 – 2922.
[18] T. H. Claus, F. Nyfeler, H. A. Muenkel, M. G. Burns, T. Pate, S. J. Pilkis, *Biochem. Biophys. Res. Commun.* **1984**, *125*, 655 – 661.
[19] M. Grötschel, L. Lovasz, A. Schrijver, *Geometric Algorithms and Combinatorial Optimisation*, 2nd revised ed, Springer, Berlin, **1993**.
[20] R. Heinrich, S. Schuster, H. G. Holzhütter, *Eur. J. Biochem.* **1991**, *201*, 1 – 21.
[21] R. Goldberg, Y. Tewari, M. Tung, *NIST Standard Reference Database 74: Thermodynamics of Enzyme-Catalyzed Reactions*, Biotechnology Division, National Institute of Standards and Technology Gaithersburg, http://wwwbiotech.nist.gov:8030/enzyme/.
[22] J. S. Van Dien, M. E. Lidstrom, *Biotechnol. Bioeng.* **2002**, *78*, 296 – 312.
[23] S. J. Van Dien, Y. Okubo, M. T. Hough, N. Korotkova, T. Taitano, M. E. Lidstrom, *Microbiology* **2003**, *149*, 601 – 619.
[24] A. Wagner, *Nat. Genet.* **2000**, *24*, 355 – 361.
[25] T. Kitami, J. H. Nadeau, *Nat. Genet.* **2002**, *32*, 191 – 194.
[26] C. C. Guet, M. B. Elowitz, W. Hsing, S. Leibler, *Science* **2002**, *296*, 1466 – 1470.
[27] T. Ideker, V. Thorsson, J. A. Ranish, R. Christmas, J. Buhler, J. K. Eng, R. Bumgarner, D. R. Goodlett, R. Aebersold, L. Hood, *Science* **2001**, *292*, 929 – 934.
[28] K. Eschrich, W. Schellenberger, E. Hofmann, *Acta Biol. Med. Ger.* **1982**, *41*, 415 – 424.
[29] D. Rother, C. G. Friedrich, *Biochim. Biophys. Acta* **2002**, *1598*, 65 – 73.
[30] C. H. Schilling, D. Letscher, B. O. Palsson, *J. Theor. Biol.* **2000**, *203*, 229 – 248.
[31] T. Dandekar, F. Moldenhauer, S. Bulik, H. Bertram, S. Schuster, *BioSystems* **2003**, *70*, 255 – 270.
[32] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, E. D. Gilles, *Nature* **2002**, *420*, 190 – 193.

[33] A. H. Tong, M. Evangelista, A. B. Parsons, H. Xu, G. D. Bader, N. Page, M. Robinson, S. Raghibizadeh, C. W. Hogue, H. Bussey, B. Andrews, M. Tyers, C. Boone, *Science* **2001**, *294*, 2364 – 2368.
[34] S. Schuster, T. Dandekar, D. A. Fell, *Trends Biotechnol.* **1999**, *17*, 53 – 60.
[35] N. D. Price, J. A. Papin, B. O. Palsson, *Genome Res.* **2002**, *12*, 760 – 769.
[36] N. D. Price, J. L. Reed, J. A. Papin, I. Famili, B. O. Palsson, *Biophys. J.* **2003**, *84*, 794 – 804.
[37] J. A. Papin, B. O. Palsson, *Biophys. J.* **2004**, *87*, 37 – 46.
[38] J. A. Papin, N. D. Price, J. S. Edwards, B. B. Palsson, *J. Theor. Biol.* **2002**, *215*, 67 – 82.
[39] S. J. Wiback, B. O. Palsson, *Biophys. J.* **2002**, *83*, 808 – 818.
[40] R. Heinrich, H. G. Holzhütter, S. Schuster, *Bull. Math. Biol.* **1987**, *49*, 539 – 595.
[41] T. Pfeiffer, S. Schuster, S. C. Bonhoeffer, *Science* **2001**, *292*, 504 – 507.
[42] B. F. Dibrov, A. M. Zhabotinsky, B. N. Kholodenko, *Bull. Math. Biol.* **1982**, *15*, 51 – 65.
[43] R. A. Majewski, M. M. Domach, *BioSystems* **1985**, *18*, 15 – 22.
[44] M. Savageau, *Biochemical Systems Analysis*, Addison-Wesley, Reading, MA, **1976**.
[45] S. Schuster, R. Heinrich, *J. Theor. Biol.* **1987**, *129*, 189–209.
[46] J. S. Edwards, B. O. Palsson, *BMC Bioinf.* **2000**, *1*, 1.
[47] P. Calik, A. Akbay, *Med. Hypotheses* **2000**, *55*, 5 – 14.
[48] C. H. Schilling, J. S. Edwards, D. Letscher, B. O. Palsson, *Biotechnol. Bioeng.* **2000**, *71*, 286 – 306.
[49] M. W. Covert, C. H. Schilling, B. Palsson, *J. Theor. Biol.* **2001**, *213*, 73 – 88.
[50] M. Giugliano, S. Cernuschi, M. Grosso, E. Aloigi, R. Miglio, *Chemosphere* **2001**, *43*, 743 – 750.
[51] R. Ramakrishna, J. S. Edwards, A. McCulloch, B. O. Palsson, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **2001**, *280*, R695 – 704.
[52] E. Svanberg, A. C. Moller-Loswick, D. E. Matthews, U. Korner, K. Lundholm, *Clin. Physiol.* **2001**, *21*, 478 – 489.
[53] J. S. Edwards, M. Covert, B. Palsson, *Environ Microbiol.* **2002**, *4*, 133 – 140.
[54] R. Mahadevan, J. S. Edwards, F. J. Doyle III, *Biophys. J.* **2002**, *83*, 1331 – 1340.
[55] K. J. Kauffman, P. Prakash, J. S. Edwards, *Curr. Opin. Biotechnol.* **2003**, *14*, 491 – 496.
[56] A. U. Raghunathan, J. R. Perez-Correa, L. T. Bieger, *Biotechnol. Bioeng.* **2003**, *84*, 700 – 709.
[57] D. Segré, D. Vitkup, G. M. Church, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15 112 – 15 117.
[58] D. E. Atkinson, *Curr. Top. Cell. Regul.* **1969**, *1*, 29 – 43.
[59] B. Papp, C. PáL, L. D. Hurst, *Nature* **2004**, *429*, 661 – 664.
[60] Note that biochemical reactions are different if they differ in at least one reactant.

Received: April 30, 2004