Available online at www.sciencedirect.com
SCIENCE DIRECT®
Metabolic Engineering 6 (2004) 378–390
METABOLIC ENGINEERING
www.elsevier.com/locate/ymben

Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics

Diana Visser[a,*], Joachim W. Schmid[b], Klaus Mauch[b], Matthias Reuss[b], Joseph J. Heijnen[c]

[a]*PURAC, Research, Development and Technology, P.O. Box 21, Arkelsedijk 46, 4200 AA Gorinchem, The Netherlands*
[b]*Institut für Bioverfahrenstechnik, University of Stuttgart, Germany*
[c]*Kluyverlaboratory for Biotechnology, Delft University of Technology, The Netherlands*

Received 11 April 2004; accepted 24 June 2004
Available online 20 August 2004

Abstract

This paper examines the validity of the linlog approach, which was recently developed in our laboratory, by comparison of two different kinetic models for the metabolic network of *Escherichia coli*. The first model is a complete mechanistic model; the second is an approximative model in which linlog kinetics are applied. The parameters of the linlog model (elasticities) are derived from the mechanistic model.

Three different optimization cases are examined. In all cases, the objective is to calculate the enzyme levels that maximize a certain flux while keeping the total amount of enzyme constant and preventing large changes of metabolite concentrations. For an average variation of metabolite levels of 10% and individual changes of a factor 2, the predicted enzyme levels, metabolite concentrations and fluxes of both models are highly similar. This similarity holds for changes in enzyme level of a factor 4–6 and for changes in fluxes up to a factor 6.

In all three cases, the predicted optimal enzyme levels could neither have been found by intuition-based approaches, nor on basis of flux control coefficients. This demonstrates that kinetic models are essential tools in Metabolic Engineering. In this respect, the linlog approach is a valuable extension of MCA, since it allows construction of kinetic models, based on MCA parameters, that can be used for constrained optimization problems and are valid for large changes of metabolite and enzyme levels.

Keywords: Linlog approach; Kinetic modeling; Metabolic control analysis

1. Introduction

With the aid of recombinant DNA technology, it has become possible to introduce specific changes in the cellular genome. This enables the directed improvement of certain properties of microorganisms, such as the productivity, which is referred to as Metabolic Engineering (Bailey, 1991; Nielsen, 1998; Stephanopoulos et al., 1998). This is potentially a great improvement compared to earlier random mutagenesis techniques, but requires that the targets for modification are known. The complexity of pathway interaction and allosteric regulation limits the success of intuition-based approaches, which often only take an isolated part of the complete system into account. Mathematical models are required to evaluate the effects of changed enzyme levels or properties on the system as a whole, using metabolic control analysis or by a dynamic sensitivity analysis (Mauch et al., 1997). Furthermore, they can be used to calculate the optimal enzyme distribution for e.g. a maximal productivity using an optimization approach (Mendes and Kell, 1998a, b; Hatzimanikatis et al., 1996, 1998; Stephanopoulos and Simpson, 1997; Torres et al., 1997; Mauch et al., 2001; Voit, 1992).

Traditionally, kinetic metabolic models are based on mechanistic rate equations, which are derived from in

*Corresponding author. Fax: +31-183-695-709.
E-mail address: d.visser@purac.com (D. Visser).

1096-7176/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2004.07.001

Nomenclature

Enzymes

AlaSynth alanine synthesis
ALD aldolase
ChoSynth chorismate synthesis
DHAPSDHAP synthases
DipimSynth diaminopimelate synthesis
ENO enolase
G1PAT glucose-1-phosphate adenyltransferase
G3PDH glycerol-3-phosphate dehydrogenase
G6PDH glucose-6-phosphate dehydrogenase
GAPDH glyceraldehyde-3-phosphate dehydrogenase
IleSynthisoleucine synthesis
KivalSynth α-ketoisovalerate synthesis
MetSynth methionine synthesis
MurSynth mureine synthesis
PFK phosphofructokinase
PGDH 6-phosphogluconate dehydrogenase
PGI glucose-6-phosphoisomerase
PGK phosphoglycerate kinase
PGM phosphoglycerate mutase
PDH pyruvate dehydrogenase
PEPCX PEP carboxylase
PGlucoM phosphoglucomutase
PK pyruvate kinase
PTS phosphotransferase system
R5PI ribose-phosphate isomerase
RPPK ribose-phosphate pyrophosphokinase
Ru5P ribulose-phosphate epimerase
Synth1 synthesis 1
Synth2 synthesis 2
TA transaldolase
TIM triosephosphate isomerase
TKa transketolase, reaction a
TKb transketolase, reaction b
TrpSynth tryptophan synthesis

Metabolites

2PG 2-phosphoglycerate
3PG 3-phosphoglycerate
6PG 6-phosphogluconate
ACCOA acetyl-coenzyme A
ADP adenosindiphosphate
AMP adenosinmonophosphate
ATP adenosintriphosphate
DHAP dihydroxyacetonephosphate
E4P erythrose-4-phosphate
F6P fructose-6-phosphate
FBP fructose-1,6-bisphosphate
G1P glucose-1-phosphate
G6P glucose-6-phosphate
GAP glyceraldehyde 3-phosphate
GLC glucose
NAD diphosphopyridindinucleotide, oxidized
NADH diphosphopyridindinucleotide, reduced
NADP diphosphopyridindinucleotide-phosphate, oxidized
NADPH diphosphopyridindinucleotide-phosphate, reduced
PEP phosphoenolpyruvate
PYR pyruvate
RIB5P ribose-5-phosphate
RIBU5P ribulose-5-phosphate
SED7P sedoheptulose-7-phosphate
XYL5P xylulose-5-phosphate
OAA oxaloacetate

Symbols

D dilution rate
e_i level of enzyme i
e_i^0 level of enzyme i at reference state
J_i^0 flux of reaction i at reference state
$K_{i,j}$ Michaelis constant of component j for reaction i
m number of intracellular metabolites
m_c number of external components
n number of reactions
$n_{i,j}$ Hill constant of component j for reaction i
v_i rate of reaction i
vmax maximal rate
x_i^0 concentration of metabolite i at reference state

Greek letters

ε_{ij} elasticity coefficient of component j for reaction i
ρ conversion factor relating the intracellular to the extracellular volume.
μ growth rate

Matrices

$\mathbf{E_c^0}$ matrix containing the elasticity coefficients of the external components
$\mathbf{E_x^0}$ matrix containing the elasticity coefficients of the intracellular metabolites
$\mathbf{S}$ stoichiometry matrix for intracellular metabolites
$\mathbf{S_c}$ stoichiometry matrix for extracellular components

Vectors

$\mathbf{c}$	vector containing external concentrations
$\mathbf{c}^0$	vector containing external concentrations at reference state
$\mathbf{c}_{feed}$	vector containing the extracellular metabolite levels in the feed
$\mathbf{e}$	vector of enzyme levels
$\mathbf{e}^0$	vector of enzyme levels at reference state
$\mathbf{f}_{pulse}$	vector of applied perturbation
$\mathbf{J}^0$	vector of fluxes at reference statev
$\mathbf{v}$	vector of reaction rates
$\mathbf{x}$	vector of metabolite levels
$\mathbf{x}^0$	vector of metabolite levels at reference state

vitro experiments. However, due to large differences between in vivo and in vitro conditions, it is unlikely that the in vitro obtained parameters are valid in vivo (Wright and Kelly, 1981; Teusink et al., 2000, Rizzi et al., 1997; Vaseghi et al., 1999). Thus, the kinetic parameters must be adjusted, using data on in vivo metabolite levels and fluxes obtained in dynamic experiments (Rizzi et al., 1997; Chassagnole et al., 2001; Vaseghi et al., 1999). Due to the complexity of mechanistic rate equations, which often contain a considerable amount of parameters, this requires a large experimental and mathematical effort.

Alternative approaches to kinetic modeling, in which nonmechanistic rate equations are applied, have also been developed (Savageau, 1976; Westerhoff and Van Dam, 1987; Hatzimanikatis et al., 1996, 1998; Hatzimanikatis and Bailey, 1997; Nielsen, 1997; Torres et al., 1997; Visser et al., 2000). The applied equations are not as complex as mechanistic rate equations and contain less parameters. They do not aim to describe the kinetics in detail; it is assumed that an approximation of the actual kinetics suffices. The rationale behind this is that metabolic redesign does not require detailed mechanistic models because of the concept of homeostasis. Homeostasis implies that the cell keeps its intracellular metabolite levels approximately constant. This means that the steady-state metabolite concentrations of the optimized metabolic network should not differ too much from the original steady state, or unexpected and undesirable side effects might occur, due to the effects changed metabolite levels have on gene expression (Stephanopoulos and Simpson, 1997; Thomas and Fell, 1998; Fell and Thomas, 1995). In other words, the extrapolation range of the kinetic metabolic model does not need to be very large, as far as metabolite levels are concerned. This reasoning suggests that one can safely apply approximative kinetic equations instead of the detailed mechanistic ones that are valid over a wide range of concentration levels.

Recently, Visser and Heijnen (2003) suggested the linlog approach for modeling of in vivo kinetics and for metabolic redesign. In this approach, all rate equations have the same mathematical structure in which the relation between rates and enzyme levels is proportional, while for metabolite levels, a linear combination of nonlinear, logarithmic concentration terms is included. All variables are considered relative to a reference steady state; the linlog approximation is valid in the neighborhood of this state, but quite large changes of metabolite concentrations, enzyme levels and fluxes are allowed. The only parameters in the kinetic equations are elasticities and steady-state fluxes.

The linlog approach was demonstrated by application to a small but complex metabolic network. Compared to a detailed mechanistic model of this network, the performance of the linlog model was satisfactory, both for simulation of dynamic perturbations and for metabolic redesign.

The next challenge is to verify whether the linlog approach is also applicable to more realistic systems. This is the subject of the study presented in this paper. The mechanistic model for the metabolic network of *Escherichia coli*, which has been developed by Chassagnole et al. (2001), is used as a model system. This model is approximated by a linlog model. Both models are used to calculate the enzyme levels that maximize a certain flux while the total amount of enzyme and the metabolite levels are constrained, as suggested by Mauch et al. (2001). The predictions of both models are compared to check the validity of the linlog approach in optimization problems.

2. Theoretical aspects

2.1. The model system

The system that is studied in this paper is the dynamic model of primary metabolism of *Escherichia coli*, which has been developed by Chassagnole et al. (2001). This model was developed to gain understanding of the flux distribution, regulation phenomena and control properties in the *Escherichia coli* metabolic network. The metabolic network consists of the glycolytic pathway, including the PTS system; the pentose phosphate pathway and several branches that lead to the formation of biomass (Fig. 1). The total number of reactions (n) is 30. The 25 (m) metabolites that are included in the model are given in Table 1. The complete stoichiometry of the network can be found in Chassagnole et al. (2001).

The dynamic model of this metabolic system in a chemostat is given by the mass balances for the

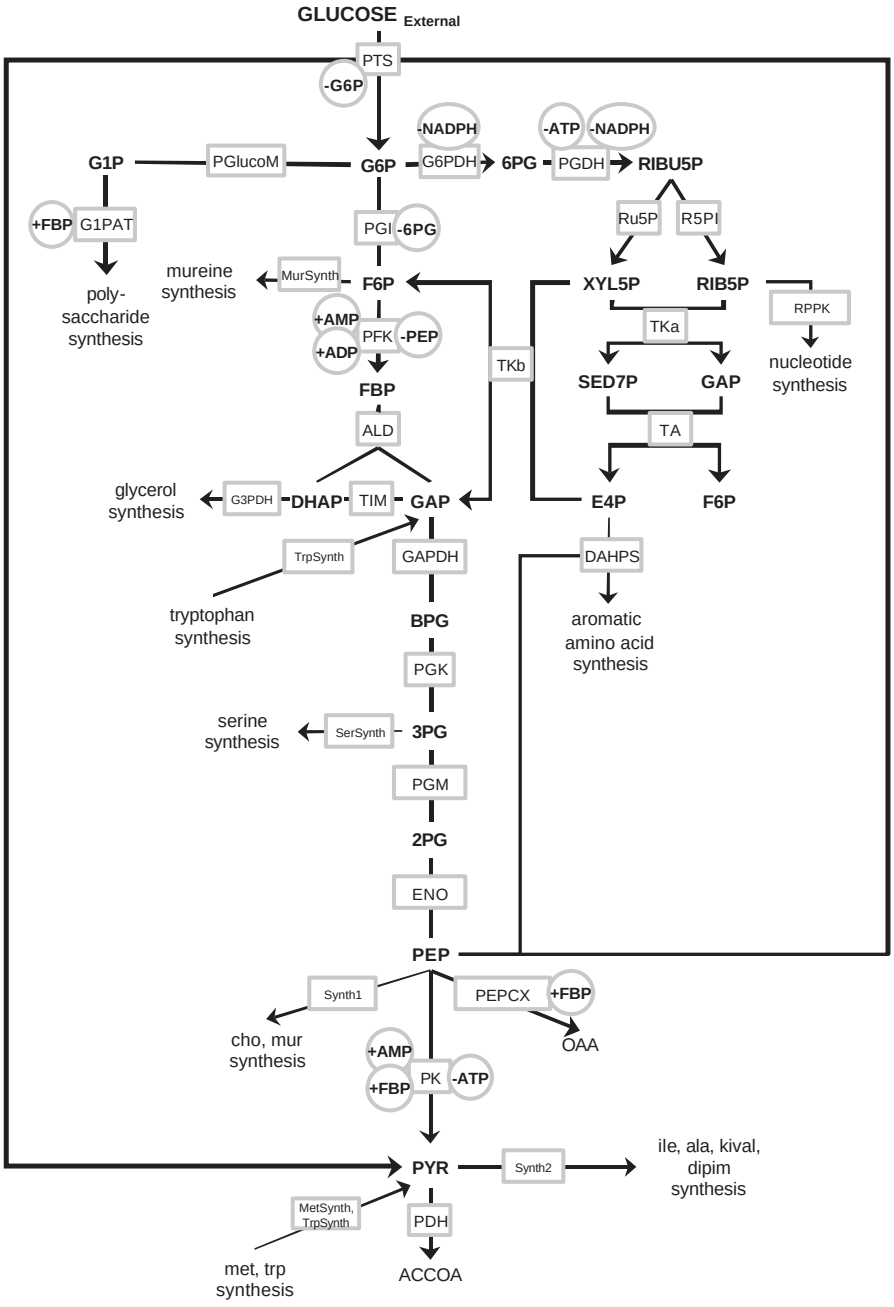

Fig. 1. Metabolic network of the central metabolism of *Escherichia coli* including glycolysis, the pentose phosphate pathway and branches to biomass formation. Allosteric effectors indicated in circles; enzymes indicated in rectangles; balanced metabolites indicated in bold.

intracellular (x) and the extracellular (c) metabolites.

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{S}\mathbf{v} - \mu\mathbf{x} \tag{1}$$

$$\frac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} = D \cdot (\mathbf{c}_{\mathbf{feed}} - \mathbf{c}) + \frac{1}{\rho} \cdot \mathbf{S}_{\mathrm{c}}\mathbf{v} + \mathbf{f}_{\mathbf{pulse}}, \tag{2}$$

$\mathbf{S}$ and $\mathbf{S}_{\mathbf{c}}$ are the $(m \times n)$ and $(m_{\mathrm{c}} \times n)$ stoichiometry matrices for the intra- and extracellular metabolites; $\mathbf{v}$ is the $(n \times 1)$ rate vector; $\mathbf{c}_{\mathbf{feed}}$ is an $(m_{\mathrm{c}} \times 1)$ vector containing the extracellular metabolite levels in the feed; D is the dilution rate; μ the growth rate and ρ is a conversion factor that relates the intracellular volume to the extracellular volume. Vector $\mathbf{f}_{\mathbf{pulse}}$ $(m_{\mathrm{c}} \times 1)$ accounts for the perturbation that is applied to identify the parameter values.

It should be noted that the model of Chassagnole et al. (2001) does not include balance equations for the nucleotides. The concentrations of these components are given by time dependent analytical functions that describe the experimentally determined response.

Table 1
Steady-state fluxes and metabolite levels (reference state)

Metabolite	Concentration (mM)	Reaction	Flux (mM/s)
Glucose	0.0556	GAPDH	0.320
G6P	3.48	PTS	0.200
F6P	0.60	PGI	0.0584
FBP	0.272	ALDO	0.141
GAP	0.218	PFK	0.141
DHAP	0.167	PGM	0.302
BPG	0.008	PGK	0.320
3PG	2.13	TIS	0.139
2PG	0.399	ENO	0.302
PEP	2.67	PK	0.0381
PYR	2.67	PEPCX	0.0427
6PG	0.808	PDH	0.188
RIBU5P	0.111	PGlucoM	0.00233
XYL5P	0.138	G1PAT	0.00232
SED7P	0.276	G3PDH	0.00212
RIB5P	0.398	SerSynth	0.0175
E4P	0.098	MurSynth	4.37E-04
G1P	0.653	DAHPS	0.00688
AMP	0.955	TrpSynth	0.00104
ADP	0.595	MetSynth	0.00226
ATP	4.27	Synth1	0.0142
NADP	0.195	Synth2	0.0536
NADPH	0.062	G6PDH	0.139
NAD	1.47	PGDH	0.139
NADH	0.1	Ru5P	0.0838
		R5PI	0.0556
		TKa	0.0453
		TKb	0.0384
		TA	0.0453
		RPPK	0.0103

2.2. The mechanistic model

The original kinetic model of the *Escherichia coli* network uses highly nonlinear mechanistic rate equations for most of the reactions. Details of these equations are given in Chassagnole et al. (2001). Table 2 gives an overview of the different types of kinetics applied. A list of the reactants and allosteric effectors for each reaction is given in Table 3. The model parameters were estimated from dynamic glucose pulse experiments. The identified parameter values can be found in Chassagnole et al. (2001). The obtained model describes the observed dynamics during the glucose pulse experiment reasonably well.

2.3. The linlog model

The kinetic equations of the linlog model are given by the following general equation:

$$\frac{\mathbf{v}}{\mathbf{J^0}} = \left[\frac{\mathbf{e}}{\mathbf{e^0}}\right]\left(\mathbf{i} + \mathbf{E_x^0}\cdot \ln\left(\frac{\mathbf{x}}{\mathbf{x^0}}\right) + \mathbf{E_c^0}\cdot \ln\left(\frac{\mathbf{c}}{\mathbf{c^0}}\right)\right), \quad (3)$$

$\mathbf{v}/\mathbf{J^0}$ is a vector of size n with elements v_i/J_i^0; $\left[\frac{\mathbf{e}}{\mathbf{e^0}}\right]$ is a square diagonal matrix containing the relative enzyme levels; i is a unit vector of size n; $\mathbf{E_x^0}$ and $\mathbf{E_c^0}$ are the $(n\times m)$ and $(n\times m_c)$ elasticity matrices; $\ln(\mathbf{x}/\mathbf{x^0})$ is a vector of size m with elements $\ln(x_i/x_i^0)$; $\ln(\mathbf{c}/\mathbf{c^0})$ is a vector of size m_c with elements $\ln(c_i/c_i^0)$.

As can be seen in Eq. (3), the variables are defined relative to a reference steady state, with metabolite levels $\mathbf{x^0}$, fluxes $\mathbf{J^0}$, enzyme levels $\mathbf{e^0}$ and extracellular concentrations $\mathbf{c^0}$. In this case, the reference steady state is the experimentally determined steady state (Table 1). The values of the elasticities (Table 3) are derived analytically from the mechanistic model:

$$\varepsilon_{ij}^0 = \frac{\mathbf{x}_j^0}{J_i^0}\left(\frac{\partial v_i}{\partial x_j}\right)^0. \quad (4)$$

The total number of parameters (elasticities) is 65, of which 17 are not independent due to the presence of near-equilibrium reactions (Appendix A). In practice, when a detailed mechanistic model is not available, one would derive the elasticities directly from metabolite and flux measurements. This can be done either by traditional MCA methods (see e.g. Fell, 1992) or by fitting of the model to experimental data (Rizzi et al., 1997; Chassagnole et al., 2001; Vaseghi et al., 1999).

Note that the reaction rate is proportional to the relative enzyme level. Hatzimanikatis (Hatzimanikatis et al., 1996, 1998; Hatzimanikatis and Bailey, 1997) suggested a similar relation in which the effect of enzyme levels on the reaction rate is described using a linear approximation. The proportionality used in the linlog format ensures that the equation is valid for large changes of the enzyme level. The extrapolation range of the format suggested by Hatzimanikatis is much smaller.

2.4. Optimization studies

The validity of the linlog approximation is examined by calculation of the optimal enzyme levels for three different cases. Optimal enzyme levels are obtained by non-linear constrained optimization (Mauch et al., 2001), using a gradient method. The gradients are derived as forward differences, for which the enzyme levels are perturbed relative to the steady-state enzyme level by $10^{-5}e^0$ and the resulting changes in the steady-state fluxes are determined by numerical simulation. In theory, this method can converge to a local optimum, but this has not been observed in our simulation studies. Calculations are carried out both for the full mechanistic model and for the linlog model. The predictions of both models are compared to identify possible deviations. In all three cases, the optimization objective is to maximize a certain flux, while keeping the total enzyme level constant (constraint 1) and preventing large changes of metabolite concentrations (constraint 2). Numerical simulation is used to verify the adherence to the constraints during the optimization process and to

Table 2
Overview of kinetics used in mechanistic model

Reaction		No. of parameters
PTS	Based on PTS rate equation Liao et al. (1996) with G6P inhibition (Kaback, 1969; Clark and Holms, 1976)	5
PGI	Reversible Michaelis Menten kinetics (Richter et al., 1975) with 6PG inhibition (Schreyer and Bock, 1980)	5
PFK	Four-state allosteric model with AMP and ADP activation (Hofmann and Kopperschläger, 1982) with PEP inhibition (Kotlarz et al., 1975) instead of ATP inhibition	9
ALD	Ordered uni–bi mechanism	6
TIM	Reversible Michaelis Menten kinetics (Richter et al., 1975)	3
GAPDH	Two substrate reversible Michaelis Menten	5
PGK	Two substrate reversible Michaelis Menten kinetics	5
PGM	Reversible Michaelis Menten kinetics	3
ENO	Reversible Michaelis Menten kinetics	3
PK	Based on equation of Johannes and Hess (1973) with additional term for AMP activation	7
PDH	Empirical Hill equation	2
PEPCX	Empirical equation including allosteric FBP activation (Kameshita et al., 1979; Yoshinaga, 1977)	3
G6PDH	Based on equation of Vaseghi et al. (1999), ATP inhibtion removed (Sanwal, 1970)	4
PGDH	Based on equation of Vaseghi et al. (1999), ATP and NADPH inhibition similar (Orozco de Silva, 1979)	4
Ru5P	First order kinetics (Vaseghi et al., 1999)	1[a]
R5PI	First order kinetics (Vaseghi et al., 1999)	1[a]
TKa	First order kinetics (Vaseghi et al., 1999)	2[a]
TKb	First order kinetics (Vaseghi et al., 1999)	2[a]
TA	First order kinetics (Vaseghi et al., 1999)	2[a]
RPPK	Michaelis Menten kinetics	1
DAHPS	Two substrates Hill equation (Akowski and Bauerle, 1997)	4
PGlucoM	Reversible Michaelis Menten kinetics	3
G1PAT	Empirical two-substrate equation including allosteric FBP activation (Preiss et al., 1975)	4
G3PDH	Michaelis Menten kinetics	1
SerSynth	Michaelis Menten kinetics	1
MurSynth	Constant	0
TrpSynth	Constant	0
MetSynth	Constant	0
Synth1	Empirical equation	1
Synth2	Empirical equation	1
Total		88

Note that some of these kinetic equations deviate from the model published by Chassagnole et al. (2001).
[a]Refers to power of metabolite concentration in equation, which is assumed to equal 1.

check whether the optimized state is stable. It should be noted that the levels of the unbalanced metabolites (the nucleotides) are assumed not to change. This assumption is necessary, since the mechanistic model (and therefore also the linlog model) does not include differential equations for these components.

The first constraint is imposed to maintain the total cell protein level at a constant level. Cell protein typically occupies 20–30% of the cell volume and it can be argued that this amount is the maximum compatible with cell function. Increased protein levels could cause diffusion problems, due to increased viscosity of the cytoplasm. Furthermore, the solubility limits of certain proteins could be exceeded, which would cause salting out of these proteins (Brown, 1991). High protein levels might also cause conformation based, secondary kinetic effects (crowding). Finally, a large increase of the protein level would require an increased production of amino acids and ATP, which might cause stress (Mauch et al., 2001).

The second constraint concerns the principle of homeostasis, which implies that the cell tries to maintain intracellular metabolites at a more or less constant level. If an optimized metabolic system would give rise to very different metabolite levels, it can be expected that these levels would influence gene expression in some way to return the system to its original state. Since the mechanisms of gene expression are usually not included in metabolic models, these effects cannot be foreseen. Additionally, increased metabolite levels can have secondary effects, such as aspecific enzyme inhibition and chemical degradation of metabolites. It is therefore desirable to look for an optimum that satisfies the homeostatic constraint (Stephanopoulos and Simpson, 1997; Thomas and Fell, 1998; Fell and Thomas, 1995; Mauch et al., 2001).

Table 3
Values of the elasticities for the reactants and allosteric effectors at the reference state given in Table 1

Reaction		No. of parameters
PTS	ε(PYR) = −1, ε(G6P) = −3.6, ε(Glc) = 1, ε(PEP) = 1, ε(Glc) = 1	5
PGI	ε(F6P) = −1996, ε(6PG) = −0.55, ε(G6P) = 1996	3 (1)
PFK	ε(PEP) = −2, ε(F6P) = 5.7, ε(AMP) = −0.48, ε(ADP) = −0.36, ε(ATP) = 0.032	5
ALD	ε(FBP) = 14, ε(GAP) = −13, ε(DHAP) = −13	3 (0)
TIM	ε(GAP) = −16, ε(DHAP) = 16	2 (0)
GAPDH	ε(GAP) = 1, ε(BPG) = −1, ε(NAD) = 0.16, ε(NADH) = −0.017	4
PGK	ε(BPG) = 85, ε(3PG) = −85, ε(ADP) = 85, ε(ATP) = −85	4 (0)
PGM	ε(2PG) = −245, ε(3PG) = 245	2 (0)
ENO	ε(2PG) = 175, ε(PEP) = −175	2 (0)
PK	ε(FBP) = 6.9E-05, ε(PEP) = 0.10, ε(ADP) = 0.30, ε(AMP) = 0.00023, ε(ATP) = −5.6E-05	5
PDH	ε(PYR) = 3.6	1
PEPCX	ε(PEP) = 0.52, ε(FBP) = 0.066	2
G6PDH	ε(G6P) = 0.81, ε(NADP) = 0.48, ε(NADPH) = −0.42	3
PGDH	ε(6PG) = 0.98, ε(NADP) = 0.59, ε(ATP) = −0.012, ε(NADPH) = −0.48	4
Ru5P	ε(RIBU5P) = 1	1
R5PI	ε(RIBU5P) = 1	1
TKa	ε(XYL5P) = 1, ε(RIB5P) = 1	2
TKb	ε(XYL5P) = 1, ε(E4P) = 1	2
TA	ε(SED7P) = 1, ε(GAP) = 1	2
RPPK	ε(RIB5P) = 0.2	1
DAHPS	ε(PEP) = 0.0013, ε(E4P) = 2.4	2
PGlucoM	ε(G6P) = 23, ε(G1P) = −23	2 (0)
G1PAT	ε(FBP) = 0.88, ε(G1P) = 0.83, ε(ATP) = 0.51	3
G3PDH	ε(DHAP) = 0.86	1
SerSynth	ε(3PG) = 0.32	1
MurSynth		0
TrpSynth		0
MetSynth		0
Synth1	ε(PEP) = 0.27	1
Synth2	ε(PYR) = 0.27	1
Total		65 (48)

In case of near-equilibrium reactions, the number of independent elasticities is given between brackets.

3. Results and discussion

3.1. Case 1

The first optimization problem concerns the glucose uptake rate through PTS, which should be maximized by changing the levels of the glycolytic enzymes (PTS, PGI, PFK, ALD, TIM, GAPDH, PGK, PGM, ENOL, PK). For clarity of presentation, all other enzymes are kept constant. In the mechanistic model, changes of enzyme levels are represented as changes in the maximal reaction rates. The optimization is restricted by the following constraints:

$$\sum \frac{\mathbf{e}}{\mathbf{e}_0} = n \tag{5}$$

$$\frac{1}{m}\sum_m \frac{|x_i - x_i^0|}{x_i^0} < 10\%. \tag{6}$$

The enzyme levels in Eq. (5) are given relative to their reference values, because the enzyme levels at the original wild type steady state are not known. Absolute values are thus not taken into account; however, Eq. (5) assures that an increase in one enzyme is compensated by a decrease of another enzyme. This constraint is applied to all of the enzymes that are being optimized. Eq. (6) states that the average metabolite levels should not change by more than 10%.

Fig. 2 shows the flux control coefficients (fcc) of the glycolytic enzymes for the PTS flux in the reference state. PTS and PFK appear to have the highest control over the PTS flux.

The optimal enzyme levels for a maximal glucose uptake rate that satisfy both constraints, obtained by constrained optimization, are given in Fig. 3A. Besides PTS and PFK, the enzymes having high fcc's with respect to PTS, also the levels of GAPDH and PK have increased. The increase of the latter is caused by the constraints that were imposed on enzyme and metabolite levels. Furthermore, Fig. 3A shows that the predictions of both models are similar. Clearly, the linlog approximation performs satisfactory.

The difference between the original and optimized state is depicted in Figs. 3B and 4. Fig. 3B shows that

the changes are negligible for most of metabolites. Only the levels of FBP, DHAP, PEP and pyruvate changed significantly. The level of FBP even increased with a factor two. As can be seen in Fig. 4, the PTS flux increased with 50%. Both for the metabolite levels and for the fluxes, the predictions of both models agree well. This is remarkable, since the difference between the reference state and the optimized state is rather large, especially for the glycolytic fluxes, for most of the enzyme levels and for the concentration of FBP. Apparently, the linlog approximation is still valid for these large changes.

It should be noted that the total number of enzyme manipulations that follows from the numerical optimization is large; the levels of all enzymes need to be adapted. In practice, it would be advisable only to manipulate those enzymes for which the model suggests a significant change. A simulation of the anticipated manipulations can be carried out to check their effect on the enzyme and metabolite levels, and the optimized fluxes.

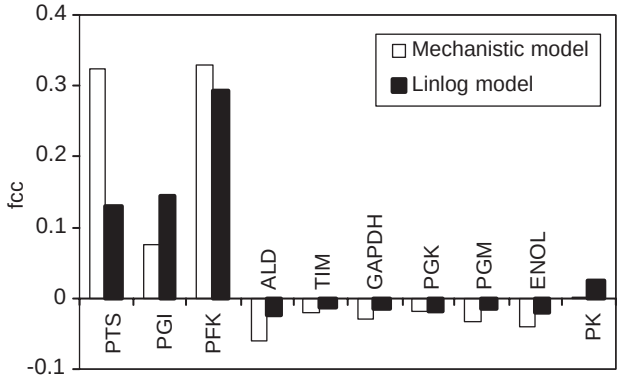

Fig. 2. The flux control coefficients of the glycolytic enzymes for the PTS flux. By definition, these should be the same for both models, since the elasticities of both models are equal. The deviations that can be observed are caused by small numerical errors due to round off in the calculation of the flux control coefficients.

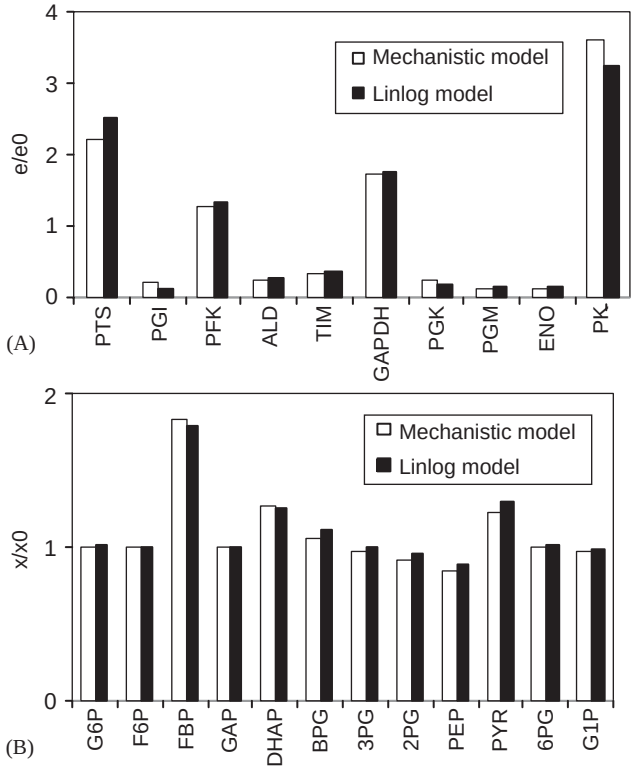

Fig. 3. Predictions for design 1. (A) Enzyme ratios. (B) Metabolite levels.

3.2. Case 2

Also in this case, the objective of the optimization is to maximize the glucose uptake rate by varying the

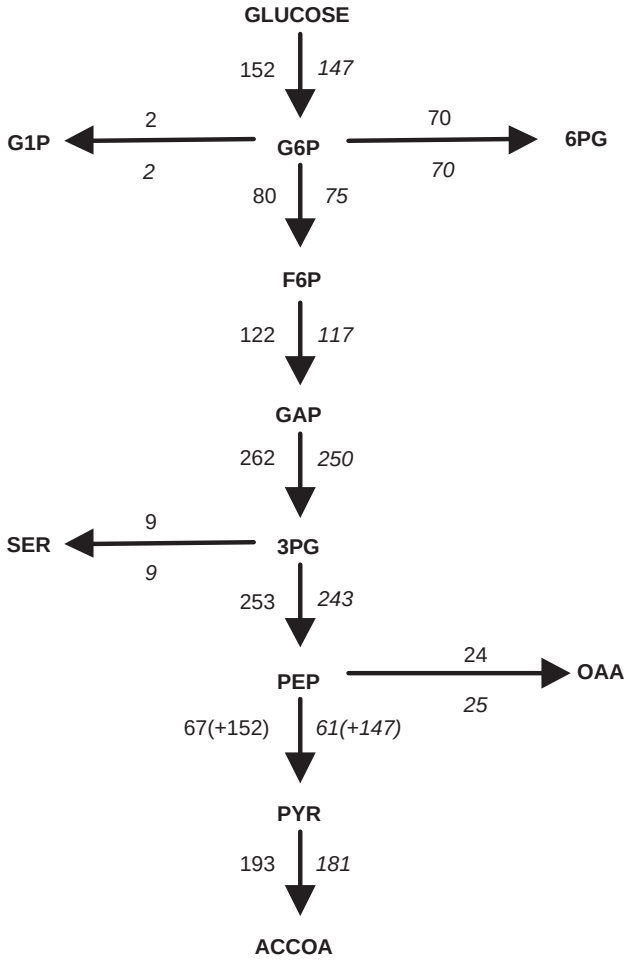

Fig. 4. Predictions for design 1. Fluxes relative to the PTS flux at the reference state. For the reaction from PEP to PYR, the amount converted via PTS is given between brackets. Italic numbers refer to the fluxes calculated by the linlog model.

levels of the glycolytic enzymes. The constraints however, are changed; constraint 1 is again given by Eq. (5), while constraint 2 is given by Eq. (7).

$$\frac{1}{m}\sum_{m}\frac{|x_i - x_i^0|}{x_i^0} < 30\%. \quad (7)$$

The allowed variation of the average metabolite level is increased from 10% to 30%. The results of the optimization with the altered concentration constraint are given in Figs. 5 and 6. In Fig. 5A can be seen that the enzyme levels predicted by the linlog model do not differ from the levels predicted for case 1. The resulting fluxes are also similar even though the concentration of FBP increased with a factor 4.3.

For the mechanistic model however, the predictions did change significantly. The main difference compared to case 1 is the predicted level of PK, which decreased with a factor 10. This decrease allowed an increase of the levels of PTS, PFK and GAPDH due to constraint 1. Furthermore, the level of FBP increased with a factor 4.3. The effect of these changes is substantial, as can be seen in Fig. 6. Compared to case 1, the PTS flux increased with another 100%, which makes the total increase 150%.

The reason of this large increase becomes clear when Figs. 4 and 6 are compared. In case 1, the efflux from

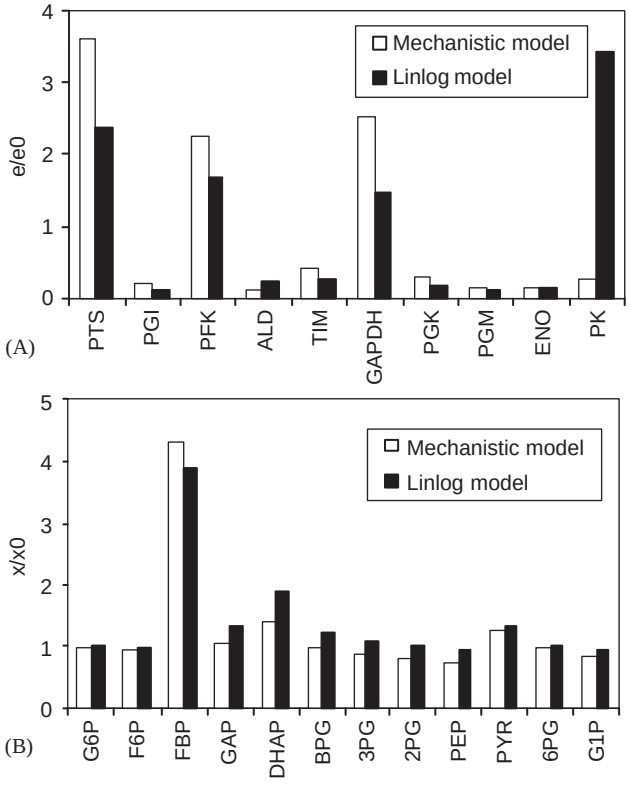

Fig. 5. Predictions for design 2. (A) Enzyme ratios. (B) Metabolite levels.

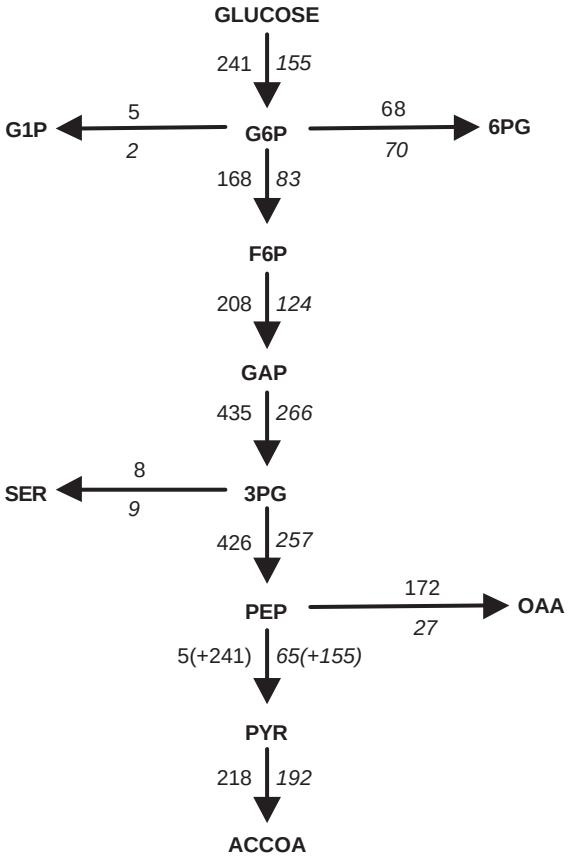

Fig. 6. Predictions for design 2. Fluxes relative to the PTS flux at the reference state. For the reaction from PEP to PYR, the amount converted via PTS is given between brackets. Italic numbers refer to the fluxes calculated by the linlog model.

glycolysis is via PK and PDH; the rate of PK has increased due to the increased enzyme level while the rate of PDH increased because of the increased concentration of pyruvate (note that the amount of PDH was not varied during the optimization). For case 2, there is a large additional efflux via PEPCX. The flux of this reaction increased with a factor 7 compared to the reference state and to case 1. This large increase is caused by the increased level of FBP, which is an allosteric activator of PEPCX. The rate equation of PEPCX is given by

$$v_{\mathrm{PEPCX}} = \frac{v\mathrm{max}_{\mathrm{PEPCX}} \cdot [\mathrm{PEP}] \cdot \left(1 + \left(\frac{[\mathrm{FBP}]}{K_{\mathrm{PEPCX,FBP}}}\right)^{n_{\mathrm{PEPCX,FBP}}}\right)}{K_{\mathrm{PEPCX,PEP}} + [\mathrm{PEP}]} \quad (8)$$

$n_{\mathrm{PEPCX,FBP}}$ equals 4.21, which means that small changes in the FBP concentration can have large effects. For the reference state, $\left(\frac{[\mathrm{FBP}]}{K_{\mathrm{PEPCX,FBP}}}\right)^{n_{\mathrm{PEPCX,FBP}}}$ equals 0.0187, which means that the contribution of FBP to the PEPCX rate

is negligible. For case 1, $\left(\frac{[FBP]}{K_{PEPCX,FBP}}\right)^{n_{PEPCX,FBP}}$ increased to 0.249, which is a factor 10 higher, but still only a small contribution to the overall rate. However, when the level of FBP increases with a factor 4.3 (case 2), the value of $\left(\frac{[FBP]}{K_{PEPCX,FBP}}\right)^{n_{PEPCX,FBP}}$ becomes 8.694 and the rate of PEPCX increases with a factor 7.

As can be seen in Figs. 5B and 6, the linlog model does not predict this large increase for the same high level of FBP. The linlog approximation of Eq. (8) is given by

$$v_{PEPCX} = J^0_{PEPCX}\left(1 + \varepsilon^0_{PEP}\ \ln\frac{[PEP]}{[PEP]^0} + \varepsilon^0_{FBP}\ \ln\frac{[FBP]}{[FBP]^0}\right). \quad (9)$$

The elasticity of FBP is very small ($\varepsilon^0_{FBP} = 0.066$), which corresponds to the observation that FBP does not contribute to the PEPCX rate in the reference state. However, due to this small value, the contribution of FBP remains negligible for much increased levels as well.

It should be realized that the observed deviation for PEPCX is an effect of the method by which the parameter values of the linlog model are calculated. The values of the elasticities are derived from the mechanistic model at the reference state (Eq. (4)), which implies that the strong allosteric effect of FBP on PEPCX is absent in the linlog model. In practice however, the elasticities would be identified from, e.g. short perturbation experiments. In these experiments, the levels of all metabolite would vary significantly and the allosteric effect of FBP would probably have been noticed and incorporated in the linlog model.

3.3. Case 3

The final example concerns the optimization of the production of serine. In this case, also the level of the SerSynth enzyme (a hypothetical enzyme that catalyzes the lumped formation of serine from 3PG and other precursors) is added to the list of enzymes to be optimized. Again, the optimization is constrained according to Eq. (5) (enzyme constraint) and Eq. (6) (metabolite constraint of 10%).

Fig. 7 depicts the values of the flux control coefficients for this case. As can be expected, SerSynth itself has a very high flux control coefficient. Again, also PTS and PFK have some control. The results of the constrained optimization are given in Figs. 8 and 9. The predictions of the mechanistic model and the linlog model are again highly similar, despite changes of metabolite levels up to a factor 2.

Fig. 8A shows that, in this case, the predicted enzyme ratios agree reasonably well with predictions based on

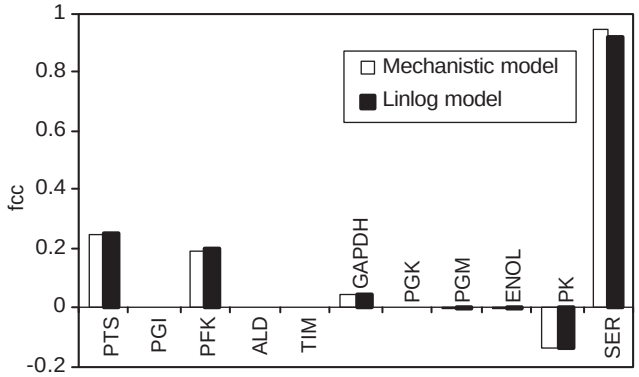

Fig. 7. The flux control coefficients of the glycolytic enzymes for the serine flux.

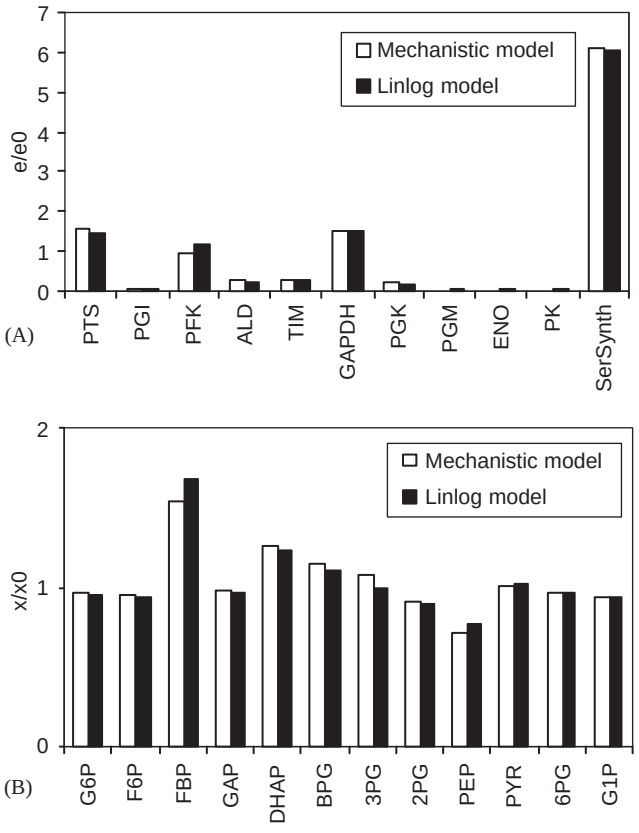

Fig. 8. Predictions for design 3. (A) Enzyme ratios. (B) Metabolite levels.

the flux control coefficients (Fig. 7). The level of SerSynth increased with a factor 6, PTS and GAPDH with a factor 1.5 and PFK remained constant. The levels of the enzymes with negligible or negative control all decreased. Fig. 8B shows that the level of FBP in the optimized state again increased considerably (50–90%). Furthermore, the levels of DHAP, BPG and 3PG increased with 10–30%. The level of PEP decreased

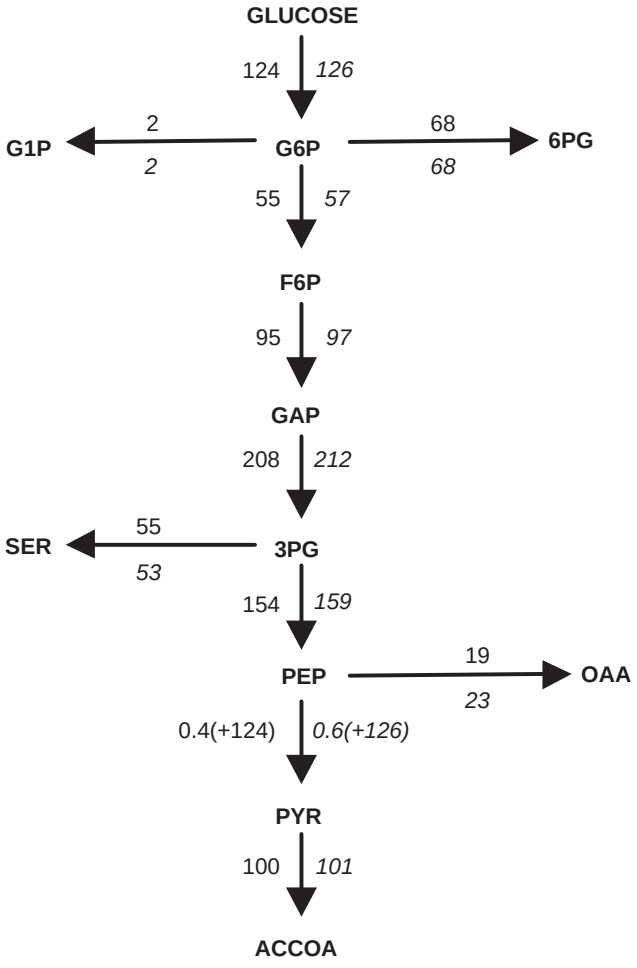

Fig. 9. Predictions for design 3. Fluxes relative to the PTS flux at the reference state. For the reaction from PEP to PYR, the amount converted via PTS is given between brackets. Italic numbers refer to the fluxes calculated by the linlog model.

with 30%. The resulting fluxes are given in Fig. 9. The serine flux increased with a factor 6. This increase is partly caused by the increased glucose uptake (25%) resulting from the increased levels of PTS, PFK and GAPDH and the increased concentration of FBP (compare cases 1 and 2). Fig. 9 shows that the flux of PK is minimized and that instead all carbon is directed towards serine. Note that still large amounts of PEP are converted into pyruvate, due to the PTS system.

4. Conclusions

The examples discussed in this paper illustrate that the linlog approach is a very useful approximation for optimization of metabolic networks. For an average variation of metabolite levels of 10% and individual changes of a factor two, a satisfactory agreement of predictions by the mechanistic model and by its linlog approximation is observed. Larger changes (factor 4) result in deviations between both models. For the cases studied, which are fairly typical, it can be concluded that a simple linlog model can approximate complete mechanistic models sufficiently well for metabolic optimization purposes.

In all examples, it would have been impossible to find the predicted optimal enzyme levels by intuitive approaches. Furthermore, the predictions did not in all cases agree with expectations based on the flux control coefficients. This demonstrates that a constrained optimization requires a complete kinetic model. In this respect, we have shown that the linlog approach is a valuable extension of metabolic control analysis, since its allows construction of kinetic models, based on MCA parameters, that can be used for constrained optimization problems and are valid for large changes of metabolite and enzyme levels.

Appendix A

Consider the near-equilibrium reactions given in Table 4. Their rates can be described by reversible mass action kinetics, where rate constant k has a large value. Using Eq. (4), the elasticities for the components involved can be derived from the rate equations. As can be seen (Table 4), the resulting elasticities are not independent; per reaction all elasticities can be calculated when only one of them is known. Furthermore, it can easily be shown that the values of the elasticities must be very high for near-equilibrium reactions, since for those reactions $K_{eq} \approx [B]^0/[A]^0$:

$$\varepsilon_A = \frac{[A]^0}{[A]^0 - [B]^0/K_{eq}} = \frac{1}{1 - \frac{[B]^0/[A]^0}{K_{eq}}}. \quad (10)$$

This implies that near-equilibrium reactions actually do not have any kinetic parameters. In practice, one can just choose high values for the elasticities of these reactions, provided that the relations between the elasticities hold (Table 4). The values of these elasticities should therefore not be considered as biologically meaningful, only the relations between them are. The system is robust when these relations are fulfilled.

For the *Escherichia coli* model, the elasticities of the near-equilibrium reactions (PGI, ALD, TIM, PGK, PGM, ENO and PGlucoM) fulfill the derived dependencies (Table 3). As can be seen, the values of these elasticities are indeed much higher than those of the non-equilibrium reactions.

Table 4
Rate equations and elasticities for several near-equilibrium reactions

Reaction	Rate equation	Elasticities	Dependencies
$A \Leftrightarrow B$	$v = k\left([A] - \frac{[B]}{K_{eq}}\right)$	$\varepsilon_A = \frac{[A]^0}{[A]^0 - [B]^0/K_{eq}}$	$\varepsilon_A + \varepsilon_B = 1$
		$\varepsilon_B = -\frac{[B]^0/K_{eq}}{[A]^0 - [B]^0/K_{eq}}$	
$A \Leftrightarrow B + C$	$v = k\left([A] - \frac{[B]\cdot[C]}{K_{eq}}\right)$	$\varepsilon_A = \frac{[A]^0}{[A]^0 - [B]^0\cdot[C]^0/K_{eq}}$	
		$\varepsilon_B = -\frac{[B]^0\cdot[C]^0/K_{eq}}{[A]^0 - [B]^0\cdot[C]^0/K_{eq}}$	$\varepsilon_A + \varepsilon_B = 1$
		$\varepsilon_C = -\frac{[B]^0\cdot[C]^0/K_{eq}}{[A]^0 - [B]^0\cdot[C]^0/K_{eq}}$	$\varepsilon_B = \varepsilon_C$
$A + B \Leftrightarrow C + D$	$v = k\left([A]\cdot[B] - \frac{[C]\cdot[D]}{K_{eq}}\right)$	$\varepsilon_A = \frac{[A]^0\cdot[B]^0}{[A]^0\cdot[B]^0 - [C]^0\cdot[D]^0/K_{eq}}$	
		$\varepsilon_B = \frac{[A]^0\cdot[B]^0}{[A]^0\cdot[B]^0 - [C]^0\cdot[D]^0/K_{eq}}$	$\varepsilon_A + \varepsilon_C = 1$
		$\varepsilon_C = -\frac{[C]^0\cdot[D]^0/K_{eq}}{[A]^0\cdot[B]^0 - [C]^0\cdot[D]^0/K_{eq}}$	$\varepsilon_A = \varepsilon_B$
		$\varepsilon_D = -\frac{[C]^0\cdot[D]^0/K_{eq}}{[A]^0\cdot[B]^0 - [C]^0\cdot[D]^0/K_{eq}}$	$\varepsilon_C = \varepsilon_D$

References

Akowski, J.P., Bauerle, R., 1997. Steady-state kinetics and inhibitor binding of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (tryptophan sensitive) from *Escherichia coli*. Biochemistry 36, 15817–15822.

Bailey, J.E., 1991. Towards a science of metabolic engineering. Science 252, 1668–1675.

Brown, G.C., 1991. Total cell protein concentration as an evolutionary constraint on the metabolic control distribution in cells. J. Theor. Biol. 153, 195–203.

Chassagnole, C., Noisommit-Rizzi, N., Schmid, J.W., Mauch, K., Reuss, M., 2001. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. Biotechnol. Bioeng. 79 (1), 53–73.

Clark, C., Holms, W.H., 1976. Control of the sequential utilization of glucose and fructose by *Escherichia coli*. J. Gen. Microbiol. 95, 191–201.

Fell, D.A., 1992. Metabolic control analysis: a survey of its theoretical and experimental development. Biochem. J. 286, 313–330.

Fell, D.A., Thomas, S., 1995. Physiological control of metabolic flux, the requirement for multisite modulation. Biochem. J. 311, 35–39.

Hatzimanikatis, V., Bailey, J.E., 1997. Effects of spatiotemporal variations on metabolic control, approximate analysis using (log) linear kinetic models. Biotechnol. Bioeng. 1, 75–87.

Hatzimanikatis, V., Floudas, C.A., Bailey, J.E., 1996. Analysis and design of metabolic reaction networks via mixed-integer linear optimization. A.I.C.H.E. J. 42, 1277–1292.

Hatzimanikatis, V., Emmerling, M., Sauer, U., Bailey, J.E., 1998. Application of mathematical tools for metabolic design of microbial ethanol production. Biotechnol. Bioeng. 58, 154–161.

Hofmann, E., Kopperschläger, G., 1982. Phosphosfructokinase in yeast. In: Wood, WA (Ed.), Methods in Enzymology. Academic Press, New York, pp. 49–60.

Johannes, K.-J., Hess, B., 1973. Allosteric kinetics of pyruvate kinase of Saccharomyces cerevisiae. J. Mol. Biol. 76, 181–205.

Kaback, H.R., 1969. Regulation of sugar transport in isolated bacterial membrane preparations from *Escherichia coli*. PNAS 63, 724–731.

Kameshita, I., Tokushige, M., Izui, K., Katsuki, H., 1979. Phosphoenolpyruvate carboxylase of *Escherichia coli*. Affinity labeling with bromopyruvate. J. Biochem. (Tokyo) 86, 1251–1257.

Kotlarz, D., Garreau, H., Buc, H., 1975. Regulation of the amount and of the activity of phosphofructokinase and pyruvate kinase in *Escherichia coli*. BBA 381, 257–268.

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

Mauch, K., Arnold, S., Reuss, M., 1997. Dynamic sensitivity analysis for metabolic systems. Chem. Eng. Sci. 52, 2589–2598.

Mauch, K., Buziol, S., Schmid, J., Reuss, M., 2001. Computer aided design of metabolic networks. In: AIChE Symposium Series. Chemical Process Control-6 Conference, Tucson, Arizona.

Mendes, P., Kell, D.B., 1998a. Non-linear optimization of biochemical pathways, applications to metabolic engineering and parameter estimation. Bioinformatics 14, 869–883.

Mendes, P., Kell, D.B., 1998b. Numerical optimization and simulation for rational metabolic engineering. In: Larsson, C., Påhlman, I.-L., Gustaffson, L. (Eds.), BioThermoKinetics in the Post Genomic Era. Chalmers Reproservice, Göteborg, pp. 345–349.

Nielsen, J., 1997. Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates. Biochem. J. 321, 133–138.

Nielsen, J., 1998. Metabolic engineering, techniques for analysis of targets for genetic manipulations. Biotechnol. Bioeng. 58, 125–132.

Orozco de Silva, A., 1979. The 6-phosphogluconate dehydrogenase reaction in *Escherichia coli*. J. Biol. Chem. 254, 10237–10242.

Preiss, J., Greenberg, E., Sabraw, A., 1975. Biosynthesis of bacterial glycogen. Kinetic studies of a glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) from a glycogen-deficient mutant of *Escherichia coli* B. J. Biol. Chem. 250, 7631–7638.

Richter, O., Betz, A., Giersch, C., 1975. The response of oscillating glycolysis in the NADH/NAD system, a comparison between experiment and a computer model. BioSystem. 7, 37–146.

Rizzi, M., Baltes, M., Theobald, U., Reuss, M., 1997. In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*, II. Mathematical model. Biotechnol. Bioeng. 55, 592–608.

Sanwal, B.D., 1970. Regulatory mechanisms involving nicotinamide adenine nucleotides as allosteric effectors. III. Control of glucose 6-phosphate dehydrogenase. J. Biol. Chem. 245, 1626–1631.

Savageau, M.A., 1976. Biochemical Systems Analysis, A Study of Function and Design in Molecular Biology, Addison-Wesley, Reading, MA.

Schreyer, R., Bock, A., 1980. Phosphoglucose isomerase from *Escherichia coli* K10, purification, properties and formation under aerobic and anaerobic condition. Arch. Microbiol. 127, 289–298.

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

Stephanopoulos, G., Simpson, T., 1997. Flux amplification in complex metabolic networks. Chem. Eng. Sci. 52, 2607–2627.

Teusink, B., Passarge, J., Reijenga, C.A., Esgalhado, E., van der Weijden, C.C., Schepper, M., Walsh, M.C., Bakker, B.M., van Dam, K., Westerhoff, H.V., Snoep, J.L., 2000. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. Eur. J. Biochem. 267, 5313–5329.

Thomas, S., Fell, D.A., 1998. The role of multiple enzyme activation in metabolic flux control. Adv. Enzyme. Regul. 38, 65–85.

Torres, N.V., Voit, E.O., Glez-Alcón, C., Rodríguez, F., 1997. An indirect optimization method for biochemical systems, description of method and application to the maximization of the rate of ethanol, glycerol, and carbohydrate production in Saccharomyces cerevisiae. Biotechnol. Bioeng. 55, 758–772.

Vaseghi, S., Baumeister, A., Rizzi, M., Reuss, M., 1999. In vivo dynamics of the pentose phosphate pathway in *Saccharomyces cerevisiae*. Metab. Eng. 1, 128–140.

Visser, D., Van der Heijden, R.T.J.M., Mauch, K., Reuss, M., Heijnen, J.J., 2000. Tendency modeling, a new approach to obtain simplified kinetic models of metabolism applied to *Saccharomyces cerevisiae*. Metab. Eng. 2, 252–275.

Visser, D., Heijnen, J.J., 2003. Dynamic simulation and metabolic redesign of a branched pathway using linlog kinetics. Metab. Eng. 5, 164–176.

Voit, E.O., 1992. Optimization in integrated biochemical systems. Biotechnol. Bioeng. 40, 572–582.

Westerhoff, H.V., Van Dam, K., 1987. Thermodynamics and Control of Biological Free-Energy Transduction. Elsevier, Amsterdam.

Wright, B.E., Kelly, P.J., 1981. Kinetic models of metabolism in intact cells, tissues, and organisms. Curr. Top. Cell. Regul. 19, 103–158.

Yoshinaga, T., 1977. Structural specificity of the allosteric inhibitor of phosphoenolpyruvatecarboxylase of *Escherichia coli*. J. Biochem. (Tokyo) 81, 665–671.