ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

In Silico Genome-Scale Reconstruction and Validation of the *Corynebacterium glutamicum* Metabolic Network

Kjeld Raunkjær Kjeldsen,[1,2] Jens Nielsen[2]

[1]Center for Microbial Biotechnology, DTU Biosys, Technical University of Denmark, DK-2800 Lyngby, Denmark

[2]Agro&Ferm A/S, Limfjordsvej 4, DK-6715 Esbjerg N, Denmark; telephone: +45-45252696; fax: +45-45884148; e-mail: jn@biocentrum.dtu.dk

Received 30 January 2008; revision received 20 May 2008; accepted 24 June 2008

Published online 23 July 2008 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.22067

ABSTRACT: A genome-scale metabolic model of the Gram-positive bacteria *Corynebacterium glutamicum* ATCC 13032 was constructed comprising 446 reactions and 411 metabolites, based on the annotated genome and available biochemical information. The network was analyzed using constraint based methods. The model was extensively validated against published flux data, and flux distribution values were found to correlate well between simulations and experiments. The split pathway of the lysine synthesis pathway of *C. glutamicum* was investigated, and it was found that the direct dehydrogenase variant gave a higher lysine yield than the alternative succinyl pathway at high lysine production rates. The NADPH demand of the network was not found to be critical for lysine production until lysine yields exceeded 55% (mmol lysine (mmol glucose)$^{-1}$). The model was validated during growth on the organic acids acetate and lactate. Comparable flux values between in silico model and experimental values were seen, although some differences in the phenotypic behavior between the model and the experimental data were observed.

Biotechnol. Bioeng. 2009;102: 583–597.

KEYWORDS: *Corynebacterium glutamicum*; genome scale metabolic model; L-lysine production

Introduction

The Gram-positive bacterium *Corynebacterium glutamicum* is used for the industrial production of different amino acids, of which L-lysine and L-glutamate are produced in the largest quantities with annual production levels of 800,000 and 1,300,000 ton, respectively. Due to its commercial importance, this organism has received at lot of attention, and significant resources have been invested in the development of efficient producer strains. Various approaches have been pursued in order to improve product yield and productivity. Mutagenesis and screening have been used with success, but with the development of molecular biological methods metabolic engineering has been applied extensively also (Cremer et al., 1991; Eggeling et al., 1998). Furthermore, fluxomics (Wittmann and Heinzle, 2002), metabolomics (Strelkov et al., 2004) or a combination of methods (Krömer et al., 2004) have been used to characterize engineered strains. These studies, often where different parts of the cell have been studied separately, have given a lot of insight into *C. glutamicum*. However, with the development of high-throughput technologies it is believed that a more holistic understanding of the whole system is essential to help extracting knowledge from these data (Palsson, 2000), consequently allowing a better identification of targets improving yields and productivity. The final goal in this respect is to develop an in silico model combining kinetic information about specific reactions. Due to the complexity of biology and the lack of kinetic information, such a model has not yet been constructed for *C. glutamicum* or any other organism.

The first step towards a complete model of an organism is a genome-scale metabolic model where the annotated genome is used in combination with available experimental data to create a list of reactions that then forms the basis for a stoichiometric model. This type of models have already been constructed for a number of species such as *Saccharomyces cerevisiae* (Förster et al., 2003), *Escherichia coli* (Reed et al., 2003), *Lactococcus lactis* (Oliveira et al., 2005), *Staphylococcus aurerus* (Heinemann et al., 2005), *Streptomyces coelicolor* (Borodina et al., 2005), *Helicobacter pylori K* (Schilling et al., 2002), *Haemophilus influenzae* (Schilling and Palsson, 2000), *Methanosarcina barkeri* (Feist et al., 2006), and *Lactobacillus plantarum* (Teusink et al., 2006).

Correspondence to: J. Nielsen

Additional Supporting Information may be found in the online version of this article.

Besides from a convenient overview of the organism and its capabilities, stoichiometric models can in some cases be used to predict phenotypic behavior during different environmental and genetic conditions (Edwards and Palsson, 2000; Edwards et al., 2001; Oliveira et al., 2005), and can directly be used to test biological hypotheses (Patil et al., 2004). Stoichiometric genome-scale models can be combined with data from high-throughput techniques such as transcriptomics (Åkesson et al., 2004; Covert et al., 2004) or fluxomics (Herrgård et al., 2006) and combining this with constraint-based methods (Price et al., 2003) the prediction power of the models can be improved.

In this article we present the first step towards a systematic biological model of *C. glutamicum*. A genome-scale metabolic model of the organism is constructed from the annotated genome of the wild type strain ATCC 13032, available literature and own experimental observations. The model is validated against data found in literature under different conditions such as different biomass production burdens and growth on different carbon sources.

Materials and Methods

The interconnectivity of metabolites in a network of biological reactions is given by reaction equations defining the stoichiometric conversion of substrates into products for each reaction (Schilling et al., 1999). Reactions are enzymatic reactions converting a substrate into a product, or transport reactions moving metabolites between different parts of the system, intracellular, extracellular or between different compartments. Active reactions in the biological system are fluxes serving to dissipate or generate metabolites. Following the law of conservation of mass, a balance describing the reaction rate of a particular metabolite through a particular reaction can be written as (Stephanopoulos et al., 1998):

$$0 = r_{\text{met}} = S \cdot v \tag{1}$$

The stoichiometric matrix **S** is an $m \times n$ matrix where m is the number of metabolites and n is the number of reactions or fluxes taking place within the metabolic network. The vector **v** refers to the reaction rate of each individual reaction or flux in the metabolic network. Metabolic models usually also include constraints, which will lead to the definition of a solution space in which the solution to the network equation must lie (Price et al., 2003). Constraints in a model are dealt with by introducing constraint equations to the metabolic network, which can assign a direction of a given reaction (reversibility or irreversibility) according to known thermodynamic constraints. These equations are typically of the form $\alpha_i \leq v_i \leq \beta_i$, where α_i and β_i are the feasible lower and upper limit of the reaction rate v_i, respectively. In practice the upper and lower limits are set to arbitrarily high values when a reaction is reversible without any regulation, whereas α_i is set to zero when a reaction is irreversible. Constraint reactions can also be used to set a maximum flux through a given reaction based on biochemical information.

The above described equation system is usually underdetermined due to the fact that the number of unknown fluxes exceeds the number of metabolites in the network, leading to a number of possible solutions, and hence, no unique solution (Bonarius et al., 1997). To cope with this linear programming/optimization can be used to maximize (or minimize) for a certain metabolite objective (e.g., growth or product formation) and seeking its maximal (or minimal) value within the stoichiometrically defined domain. This procedure is often referred to as flux-balance analysis (Palsson, 2006). In the present article flux-balance analysis is carried out using linear optimization, where the objective functions used are optimizing for either growth or lysine production. Flux-balance analysis was performed using the in-house software BioOptv4.9 employing LINDO API for linear optimization. BioOptv4.9 is available by contacting the corresponding author.

Construction of Genome-Scale Metabolic Model

Genome scale metabolic models can be constructed using data from different sources. In this work the model was constructed in three steps: (i) construction of a crude model consisting of mass balances for catabolic reactions leading to the formation of metabolites; (ii) defining the reactions involved in polymer and biomass synthesis, and energy requirements; (iii) complete the metabolic network by adding missing reactions, and to revise the reactions added in the first step with respect to physiochemical conditions (reversibility of reactions).

First step was carried out based on the published annotated genome (Kalinowski et al., 2003). The KEGG database (http://www.genome.ad.jp/kegg/) and the BioCyc database (http://www.biocyc.com/) were used in this process, because these databases present annotation of each gene linked to its function. In addition, these databases give a graphical overview of the metabolism of the organism. Data for step two were found in literature from different sources, as shown in the supplementary material. Maintenance requirements for the model was determined changing the ATP demand (ATP used for maintenance and assembling of macromolecules) for maximal biomass yield on substrate (Y_{sx}) fitted experimental values. The third step was done by reviewing biochemical literature for *C. glutamicum* or by using own unpublished observations.

The complete model including a list of references used for the re-construction is available as supplementary files.

Biomass Synthesis Equations

For genome scale metabolic models the equations defining the biomass synthesis in a genome-scale model are important. The biomass-equation consists of reactions converting single molecules into macromolecules, which are building

blocks of the biomass. For each macromolecule an equation was formulated based on literature, and energy consumption associated with the assembling reactions was also included (see Supplementary Material). The representative averaged biomass composition of a *C. glutamicum* strain can be seen in Table I. Since no data on energy requirements for polymerization of macromolecules in *C. glutamicum* could be found, these values were approximated using values for *E. coli* (Ingraham et al., 1983). The macromolecule components were lumped in a final biomass assembly reaction based on their weight fraction of the biomass. In our model we included protein, DNA, RNA and cell-wall components as macromolecules. The macromolecular composition of biomass was taken from Cocaign-Bousquet et al., (1996) whereas composition of each macromolecule was taken from different references (see Supplementary Material). This same biomass composition was used for all our simulations.

Macromolecular Composition

Data for the amino acid composition of the protein fraction in *C. glutamicum* were taken from Cocaign-Bousquet et al. (1996). The composition of the DNA was calculated based on G + C-content of the genomic sequence of *C. glutamicum* ATCC 13032 (Kalinowski et al., 2003). It was assumed that RNA consisted of 5% mRNA, 75% rRNA and 20% tRNA (molar). The nucleotide composition of mRNA was taken as for genomic DNA. The nucleotide composition of rRNA was calculated from the sequences of 16S, 23S, and 5S ribosomal RNA units. And finally tRNA composition was found from sequences of leucine and glycine transporting RNAs. Sequences were downloaded from GenBank (http://www.ncbi.nlm.nih.gov).

The chemical structure of the *C. glutamicum* cell-wall has been intensively studied and consists of a complex network of the polysaccharide peptidoglycan covalently linked to another complex polysaccharide arabinogalactan, which is further esterified with mycolic acids (Puech et al., 2001). Associated with, but not covalently linked to this fraction a number of lipids are also connected to the cell-wall, the more abundant being phospholipids and trehalose mycolates (Daffé, 2005). For simplification the cell-wall components mentioned above have been divided into individual lumped reactions which are part of the biomass equation based on their fraction.

The plasma membrane of *C. glutamicum* mainly consists of polar lipids, of which phospholipids are the dominant type. In *C. glutamicum* the major phospholipids constituents are oleic acid (18:1) and palmic acid (16:0), which contribute to more than 90% of the lipid pool of the phospholipids (Collins et al., 1982b; Hoischen and Krämer, 1990). In addition to these fatty acids, myristic acid (14:0), pentadecanoic acid (15:0), stearic acid (16:1) and tuberculostearic acid (18:0) are present in minor quantities (Collins et al., 1982b). Based on the published data of *C. glutamicum* a lumped reaction with molar ratios of each individual phospholipid was made.

The stoichiometric equation formation of peptidoglycan was described by a lumped reaction. The molar composition

Table I. Metabolites considered being required for biomass in C. glutamicum.

Protein (0.52 protein g/g DCW)	mmol/g protein
Alanine	1.268
Arginine	0.361
Aspartate	0.368
Aspargine	0.368
Cysteine	0.084
Glutamate	1.044
Glutamine	0.650
Glycine	0.671
Histidine	0.128
Isoleucine	0.359
Leucine	0.669
Lysine	0.355
Methionine	0.144
Phenylalanine	0.244
Proline	0.303
Serine	0.467
Threonine	0.519
Tryptophan	0.052
Tyrosine	0.148
Valine	0.520
ATP (polymerization energy)	38.44
DNA (0.01 g DNA/g DCW)	mmol/g DNA
dAMP	0.748
dCMP	0.871
dTMP	0.748
dGMP	0.871
ATP (polymerization energy)	4.44
RNA (0.05 g RNA/g DCW)	mmol/g RNA
AMP	0.69
GMP	1.01
CMP	0.70
UMP	0.70
ATP (polymerization energy)	1.24
Lipids (0.13 g Lipids/g DCW)	
Mycolic acids (0,102 g mycolic acids/g DCW)	mmol/g mycolic acids
Trehalose monocyrynomycolate	0.061
Trehalose dicyrynomycolate	0.043
Free mycolic acids	1.875
Phospholids (0.028 g phospholipids/g DCW)	mmol/g phospholipids
Phosphatic acid	0.033
Phosphatidylglycerol	1.133
Cardiolipin	0.007
Phosphatidyllinositol	0.042
Phosphatidyllinositol mannoside	0.065
Cell wall carbohydrates (0.19 g cell wall carbohydrate/g DCW)	
Peptidoglycan (0.095 g peptidoglycan/g DCW)	mmol/g peptidoglycan
N-acetylmuramic acid	1.026
N-acetylglucosamine	1.026
L-alanine	1.026
Diaminopimelinic acid	1.026
D-glutamate	1.026
D-alanine	2.052
ATP (polymerization energy)	5.129
Arabinogalactan (0.095 g arabinogalactan/g DCW)	mmol/g arabinogalactan
Arabinose	4.650
Galactose	0.498

was estimated based on data from different sources (Keddie and Cure, 1978; Daffé, 2005). Also the arabinogalactan fraction was described by a lumped reaction due to its complexity. Data used were taken from (Puech et al., 2001).

The mycolic acid synthetic pathway is not well characterized, although some of the genes involved have been identified (Brand et al., 2003). Based on structural considerations it has been postulated that mycolic acids are synthesized by a condensation and decarboxylation reactions (Daffé, 2005). The ''free'' mycolic acids are known to form esters with trehalose in order to form mycolates were the mycolic acids are bound within the cell wall. The majority of the mycolic acid fraction consists of the 32:2 3OH and 34:1 3OH mycolic acids (Collins et al., 1982a). Other mycylic acids identified are 32:0, 34:0, 36:1, and 36:2 (Jang et al., 1997). In general two types of mycolates are considered to be present in the *C. glutamicum* cell wall: trehalose monocorynemycolate (TMCM) and trehalose decorynemycolate (TDCM) (Daffé, 2005), both of which was included in the model. Data from Daffé (2005) was used to make a lumped reaction for the mycolic acid fraction of the cell.

Energy Requirements for Growth and Maintenance

In addition to the stoichiometry of each individual macromolecule and the overall biomass assembling reaction, the stoichiometry of the growth and non-growth associated energy connected to biomass is important in a metabolic model. In our model we used the value of 29.2 mmol ATP $(\text{g biomass})^{-1}$ for growth associated ATP demand as estimated by Cocaign-Bousquet et al. (1996) based on experimentally determined macromolecule composition of *C. glutamicum* and known anabolic pathways. ATP demand for growth associated maintenance reactions was found by fitting the biomass yield to experimental value of 0.61 g biomass $(\text{g glucose})^{-1}$ found by Cocaign-Bousquet et al. (1996). ATP requirements for growth were kept constant for biomass synthesis for all simulations, as generally done in flux-balance analysis simulations.

Simulation Methods

As mentioned above, linear optimizations were conducted using the BioOptv4.9 software. Split points in the network solutions were checked to see if the results were unique, or if other optima were possible. Results presented in this article were unique, unless else is mentioned. The NADPH generating reaction catalyzed by malic enzyme (EC 1.1.1.40) and the reaction decarboxylating oxaloacetate catalyzed by oxaloacetate decarboxylase (EC 4.1.1.13) were constrained to zero when glucose was used as carbon source based on biochemical evidence (Petersen et al., 2000).

Results and Discussion

Reconstruction of the *C. glutamicum* Metabolic Network

We constructed a genome scale model of the *C. glutamicum* metabolic network (for statistics see Table II). The constructed metabolic network consisted of 446 reactions, all of which were unique (*iso*-enzymes were removed), which represent 15% of the protein encoding genes identified in *C. glutamicum* (Kalinowski et al., 2003). Four hundred eleven metabolites were involved in the metabolic network of which 55 were involved in transport reactions, and had an internal equivalent (Table II).

Energy Requirements for Biomass Formation

Macromolecule composition of biomass was taken from Cocaign-Bousquet et al. (1996). ATP demand for biomass assembly was set to 29 mmol $(\text{g biomass})^{-1}$ (Cocaign-Bousquet et al., 1996), and ATP demand for maintenance was adjusted so the maximum biomass yield (g biomass $(\text{g glucose})^{-1}$) was 0.61 (Cocaign-Bousquet et al., 1996). The ATP demand for maintenance was found to be 19 mmol $(\text{g biomass})^{-1}$, which is higher than values earlier reported in literature (1.8–9.2 mmol ATP $(\text{g biomass})^{-1}$) (de Graaf, 2000; Varela et al., 2004). The high ATP demand for maintenance may be due to an underestimated growth associated ATP demand. Another reason may be that the composition of the respiratory chain affects the P/O-ratio and hence the ATP yield (Bott and Niebisch, 2003). When maximizing for growth the in silico model uses the most efficient oxidative phosphorylation, which may not be the case in reality.

Model Validation

Simulations for model validation were carried out according to what is described in the Simulation Methods Section. Simulations were done by maximizing for lysine production and constraining biomass production.

Table II. Statistics for the genome scale model.

Total genes in *C. glutamicum* (ORFs)	3,002
Reactions in genome scale model	446
Biochemical evidence	213
Clear function (functional annotation derived from probable homologues)	209
Tentative function (functional annotation derived from tentative homologues)	22
Putative function (added reactions to complete network)	2
Active reactions during growth on glucose	199
Metabolites in genome scale model	411
Internal metabolites	356
External metabolites (metabolites involved in transport reactions)	55

Pentose Phosphate Pathway (PPP) Fluxes

The genome-scale metabolic model was validated against published data on metabolic flux distributions during various growth- and lysine production regimes, and with different strains. The publications used for this comparison were Krömer et al. (2004), Wittmann and Heinzle (2001), Wittmann and Heinzle (2002), and Vallino and Stephanopoulos (1993). The datasets Krömer et al. (2004), Wittmann and Heinzle (2001), and Wittmann and Heinzle (2002) were all based on batch cultivations where flux distributions were determined using ^{13}C-labelled glucose, whereas Vallino and Stephanopoulos (1993) used bioreaction network analysis for estimation of flux distributions. The individual datasets for the three first references were selected to give the same lysine yields (18%) and three different stains were represented (ATCC 13287, ATCC 21253, and ATCC 21526, respectively). The fourth dataset from Vallino and Stephanopoulos (1993) was selected to have a data point at a higher lysine yield (30%). All the flux data are from balanced growth conditions, and details on how the fluxes were found are given in the original articles. In silico simulations were carried out yielding lysine conversion yields in the range of the values found in the published data, in this case being 18% (mmol lysine (mmol glucose)$^{-1}$) and 30%, respectively. In all the experiments used to validate the model the biomass yields were lower than the in silico model (in same order as above: 0.38, 0.33, 0.26, and 0.27 g biomass (g glucose)$^{-1}$)) which could partly be explained by the fact that by-products in all cases were produced, which were not the case for simulations. Moreover, stoichiometric parameters such as the P/O ratio and ATP yields and maintenance requirements may change due to different strain backgrounds and growth conditions. Although all three references of Krömer et al. (2004), Wittmann and Heinzle (2001), and Wittmann and Heinzle (2002) showed lysine yields of 18%, some variation in fluxes from reference to reference were seen (Fig. 1; three first fluxes from the top). At the level of glucose-6-phosphate where this compound is converted either towards the pentose phosphate pathway (PPP) or the glycolysis a 12% absolute difference was seen (59–71% towards PPP) between the fluxes in the experiments. The in silico model predicted a 61% flux towards PPP at this split, which correlated well with two of the experiments (Fig. 1) (Krömer et al., 2004; Wittmann and Heinzle, 2002), whereas the flux was higher (71%) for the data of Wittmann and Heinzle (2001) (Fig. 1). For the three experiments the tricarboxyic acid (TCA)-flux ranged from 46% to 68% whereas the model simulated value was 51% (Fig. 1).

When simulations were carried out fitting the data of Vallino and Stephanopoulos (1993) (Fig. 1), a higher flux through the PPP was seen for the model (84%) than it was the case for literature values (69%). Instead Vallino and Stephanopoulos (1993) saw a higher flux through the TCA (46%) when compared to the in silico model (36%). Instead the anaplerotic flux from pyruvate/phosphoenolpyruvate into oxaloacetate was higher for the model (48%) than for the data of Vallino and Stephanopoulos (1993) (41%). Although differences were seen between the published data and the in silico model, fluxes were all in the same range, and coherence between experimental data and model data could be recognized.

Data from Vallino and Stephanopoulos (1993) were taken at early exponential phase, just as lysine synthesis had initiated after the depletion of threonine, and with a high growth rate. Looking at data from the same reference for late exponential phase, the difference between model predictions and observed values were bigger (data not shown), even though lysine yields were only marginally changed. The PPP-flux decreased to 41% and the TCA-flux increased to 70% (Vallino and Stephanopoulos, 1993). The change in fluxes was thought to be due to a decrease in growth rate, which were followed by ATP excess (Vallino and Stephanopoulos, 1993). The excess ATP needed to be consumed, that is, through futile cycles, which lead to a redirection of carbon from the PPP towards the TCA cycle due to biological regulatory issues. Such changes in metabolism due to biochemical regulation cannot be predicted by the model, since it doesn't take biological and biochemical regulations into account.

Glycolysis Fluxes

Model flux values through the glycolysis were in general consistent with observed experimental values. A higher drain of carbon from glycolysis was seen in data from Krömer et al. (2004) when compared to model values and other experimental values. This can partly be explained by by-product formation, where significant amounts of extracellular glycerol and dehydroxyacetone phosphate were observed (Krömer et al., 2004).

Fluxes Through the TCA and Around the Anaplerotic Node

At the level of phospoenolpyruvate and pyruvate, *C. glutamicum* possesses a number of reactions for the interconversion between glycolytic C_3 (pyruvate and phosphoenolpyruvate) and C_4 metabolites of the TCA cycle (oxaloacetate and malate). The importance of a high anaplerotic netflux for providing sufficient amounts of oxaloacetate for lysine production has been seen in numerous experiments (Fig. 2; gray line), and this relationship was also seen for the genome scale model (Fig. 2; black line). However, it was seen that the anaplerotic netflux was higher for the in silico model than what was observed for the experimental work (Fig. 2). The reason for this is partly due to a higher TCA-flux for the experimental work than for the in silico data, which will be discussed below. Comparable anaplerotic netfluxes between experimental data and simulations were observed for the data in Figure 1. When the lysine production was low (18%) an anaplerotic netflux around 40% was seen. However, for the data of the higher lysine yields (30%) a higher anaplerotic flux for the

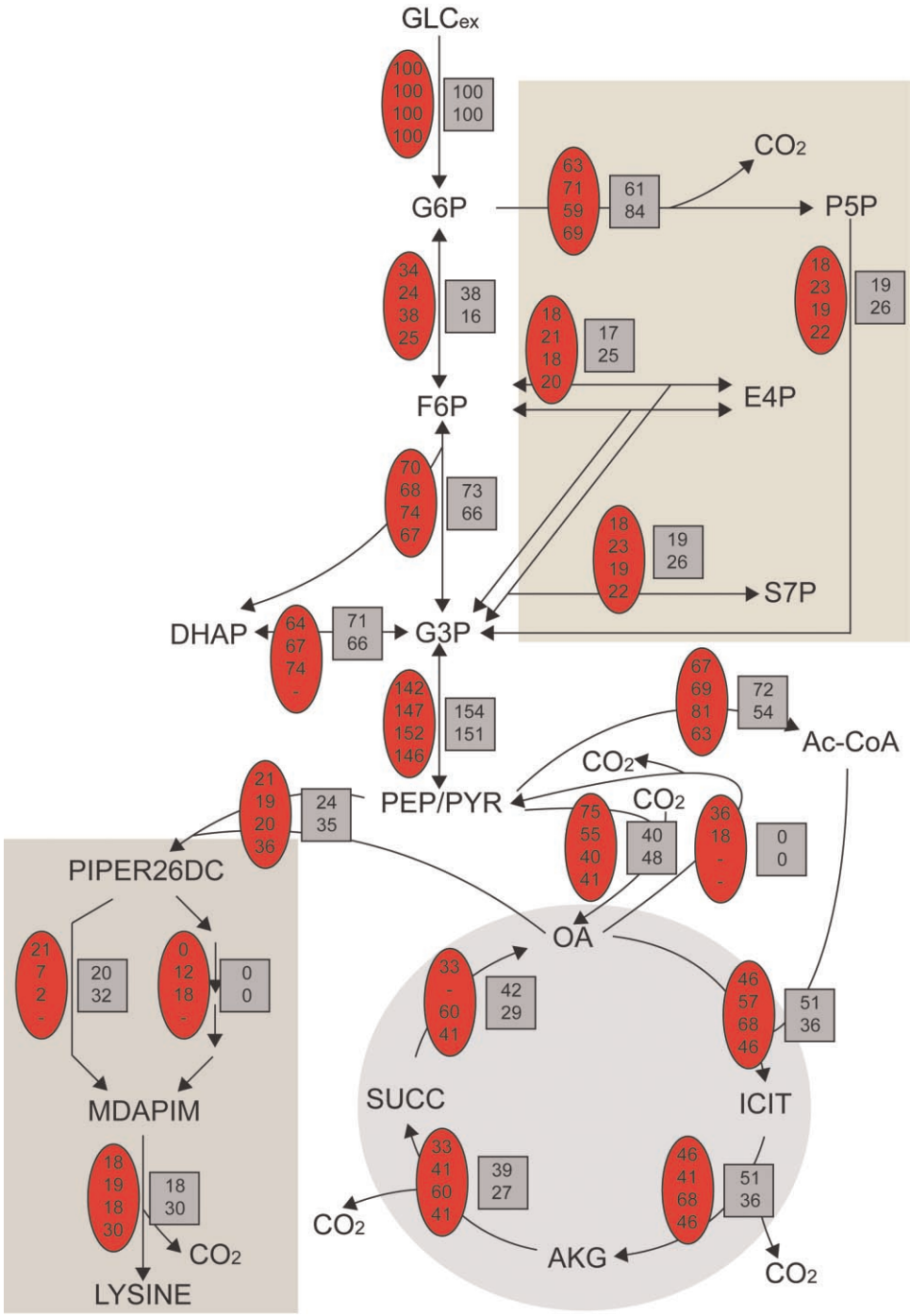

Figure 1. Flux distribution values for genome-scale model simulations (shown in gray boxes) and literature values (shown in red ovals) found by metabolic flux analysis. Flux values are expressed as molar percentage of the specific uptake of glucose. When no flux was available from the data the symbol (—) was used. Y_{sx} values for the model simulations was constrained to 0.48 and 0.40 (g biomass (g glucose)$^{-1}$), from top and down. Literature values were taken from (from top and down): Krömer et al. (2004), Wittmann and Heinzle (2001), Wittmann and Heinzle (2002), and Vallino and Stephanopoulos (1993). Abbreviations: GLC_{ex}: extracellular glucose; G6P: glucose-6-phosphate; F6P: fructose-6-phosphate; P5P: pentose-5-phosphate; E4P: erythose-4-phosphate; S7P: sedoheptulose-7-phosphate; DHAP: dihydroxyacetone phosphate; G3P: glyceraldehyde phosphate; PEP: phosphoenolpyruvate; PYR: pyruvate; Ac-CoA: acetyl CoA; ICIT: isocitrate; AKG: α-ketogluterate; SUCC: succinate; OA: oxaloacetate; PIPER26DC: L-piperidine 2,6-dicarboxylate; MDAPIM: meso-2,6-diaminopimelate. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

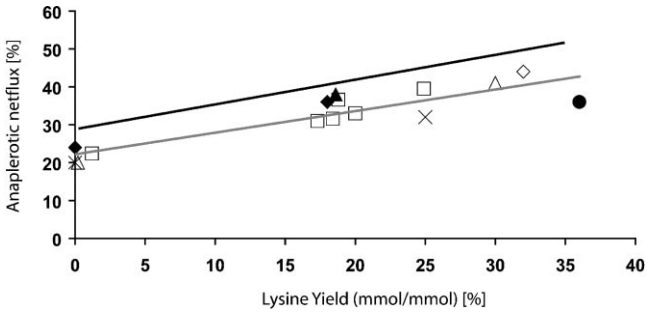

Figure 2. Anaploretic netflux at different lysine yields. Black line: Anaploretic netflux for in silico organism; Gray line: Linear regression for experimental data. White boxes: Wittmann and Heinzle (2002); White diamonds: Marx et al. (1999); White triangles: Vallino and Stephanopoulos (1993); Black circles: Vallino and Stephanopoulos (1994); Crosses: Sonntag et al. (1995); Black triangles: Marx et al. (1996); Black diamonds: Marx et al. (1997).

simulated data was seen compared to the experimental data (41% for the experiment and 48% for the model) (Fig. 1) which was in agreement with the more general observations from literature (Fig. 2).

Relationship Between PPP- and TCA-Fluxes During Lysine Production

It is known that the relationship between the TCA and the PPP is important for lysine production since a high reaction rate of the PPP is necessary to support NADPH generation when lysine synthesis is high. A tendency for this correlation can be seen from experimental data from literature when TCA- and PPP-fluxes at different lysine yields are compared (Fig. 3; symbols), a correlation which have also been suggested by Kelle et al. (2005). Such a correlation is also predicted by the model (Fig. 3; lines). In general the TCA-flux was higher for the experimental values when compared to the model, and the PPP-fluxes for the

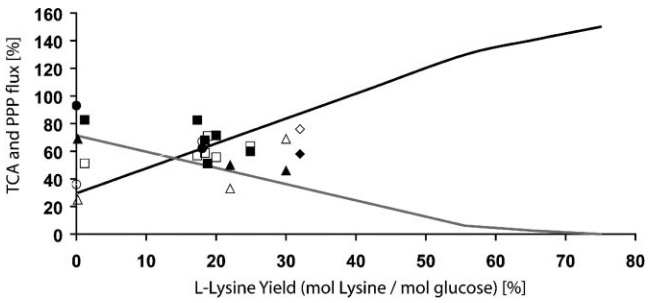

Figure 3. Flux through the TCA cycle and the Pentose phosphate pathway (PPP) at different lysine conversion yields. Black line: PPP flux for in silico organism; Gray line TCA flux for in silico organism. Open symbols are PPP fluxes and filled symbols are TCA fluxes from experimental metabolic flux analysis from different references. Squares: Wittmann and Heinzle (2002); diamonds: Marx et al. (1999); Triangles: Vallino et al. (1993); Circles: Marx et al. (1997); Crosses: PPP data from Sonntag et al. (1995).

experimental data presented here were in most cases lower than values for the corresponding in silico simulations. The reason for the generally lower TCA-fluxes for the in silico model when compared to experimental data was the higher anaplerotic netflux, which from a stoichiometric point of view is optimal for lysine production. The in silico model will per definition chose this pathway which may not be the case in reality, where biological regulation will affect the operation of the biological network.

A bend of both the TCA-flux and the PPP-flux curves were seen for the in silico flux data when lysine yield exceeded 55% (Fig. 3). At this point a flattening of the curves was observed (the TCA-flux decreased at a lower rate and the PPP-flux increased at a lower rate). This metabolic change is discussed further in the next section. Some of the differences in flux distributions between the experiments mentioned above can probably be explained by different strain backgrounds, and different cultivation conditions.

Maximizing for Lysine Production

Simulations for lysine production were done by maximizing for lysine production as the objective function, and constraining biomass production at different levels (see text for details). Simulations were carried out according to what is described in the Simulation Methods Section.

Negative Effect of ATP Excess on Lysine Production

The maximum lysine yields were found constraining the biomass yields at different levels using the genome-scale model (Figs. 4A and 5A). The model predicted a maximum lysine yield of 75% when no biomass was produced (Figs. 4A and 5A), which is the maximum theoretical value for this organism (Stephanopoulos and Vallino, 1991). As biomass synthesis increased towards the maximum, lysine yields decreased (Figs. 4A and 5A). The decrease observed was not linear as a bend of the curve was observed when the lysine yield was around 55%, and Y_{sx}-values were 20%, corresponding to a decreasing Y_{xp}-value: less lysine was produced per cell unit (Fig. 4A). Also when plotting the TCA- and PPP-flux data versus the lysine yield, a bend is observed at a yield of 55% (Fig. 3). This shift in the central metabolism was investigated, and it was found that the change was due to a change in ATP availability, where the regime was changed from an ATP limiting condition to an ATP excess condition. During the synthesis of lysine in *C. glutamicum*, a net ATP-production is seen. ATP is consumed for biomass related reactions, however when the ATP produced during lysine production exceeds the ATP demand for biomass synthesis, ATP is in excess. The way the model copes with this, is to burn ATP in a number of futile cycles, of which an example can be recognized in Figure 5A, bottom flux, where a futile cycle between the C_3-pool (pyruvate) and the C_4-pool (oxaloacetate) can be recognized with the net reaction of ATP $\rightarrow$ AMP + Pi + Pi. This change

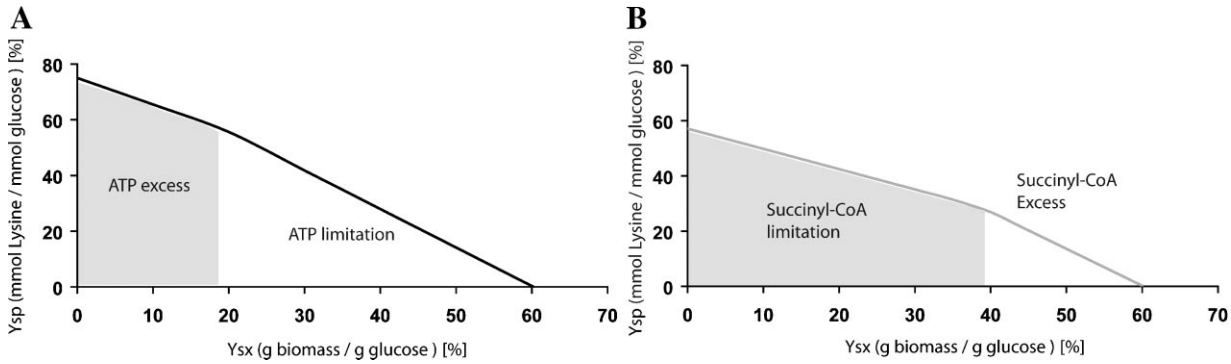

Figure 4. Simulation results for in silico model when maximizing for lysine production and constraining biomass production. Results presented as lysine yield (Y_{sp}) as a function of the biomass yield (Y_{sx}) both expressed as percentage. **A**: In silico model with dehydrogenase branch 100%. **B**: In silico model with dehydrogenase branch constrained to zero.

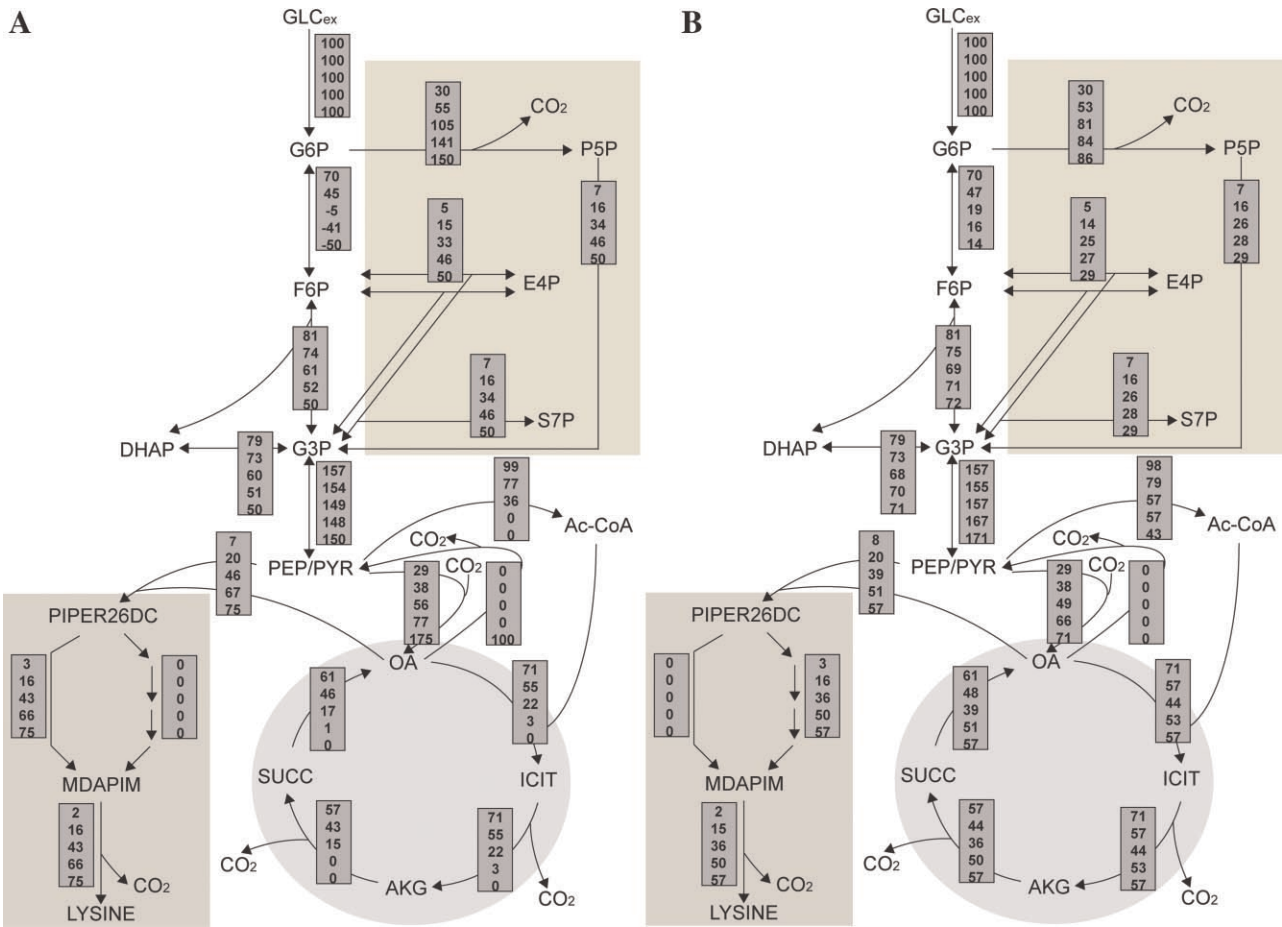

Figure 5. Flux distribution values (shown in gray boxes) for in silico simulations maximizing for lysine production constraining biomass production at different levels. Fluxes expressed as molar percentage of specific glucose uptake. Biomass yield (Y_{sx}) constrained to (from top and down) 0.60, 0.50, 0.30, 0.10, and 0.0 (g biomass (glucose)$^{-1}$. **A**: Direct dehydrogenase variant of lysine synthesis pathway used. **B**: Succinylase variant of lysine synthesis pathway used. Abbreviations see Figure 1. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

from ATP limitation to ATP excess changes the regime and hence can explain the changes observed in Figures 3 and 5A.

Split in the Lysine Synthetic Pathway and Its Effect on Lysine Yield

In the lysine synthesis pathway of C. glutamicum two pathways are possible. Either the direct dehydrogenase variant (ddh), or the succinylase variant which involves four reactions (Schrumpf et al., 1991). When using the genome-scale model for maximizing for lysine production the dehydrogenase variant is always preferred. This is due to the fact that the succinylase variant consumes the intermediate succinyl-CoA, which is an intermediate of the TCA cycle, requiring this to be active during lysine synthesis, which is not the case when the dehydrogenase variant is used. The exclusive use of the dehydrogenase variant is however far from reality based on experimental data, which shows that both variants most often are used at different ratios. Experiments have shown that the ratio between the two variants varies between a ratio higher than 2:3 for the succinylase variant with values ranging from 67–68% (Sonntag et al., 1993) to 72–89% (Wittmann and Heinzle, 2002), and with the dehydrogenase variant being the dominant pathway with values of 63% (Wittmann and Heinzle, 2001) using this pathway. The flux ratio between the two variants was even shown to change through a fermentation starting with a 72% flux through the dehydrogenase variant, decreasing to around zero at the end of the fermentation (Sonntag et al., 1993). This difference in the ratio of these two pathways was also seen for the experiments selected for the validation (Fig. 1). Either the dehydrogenase variant was used alone (Fig. 1; top), or a combination of the dehydrogenase variant and the succinyl variant with different ratios was seen (Fig. 1; 2nd and 3rd value from the top). As mentioned earlier a significant difference was seen between the TCA-fluxes for the experiments shown in Figure 1. In addition it can be seen that for the first three experiments (Krömer et al., 2004; Wittmann and Heinzle, 2001, 2002), the flux through the TCA cycle and the route which was used through the lysine pathway (dehydrogenase variant or succinylase variant) correlated, so when a high flux through the succinylase variant was seen, a higher TCA-flux could be observed. This observation may explain some of the differences seen between the TCA-fluxes and the anaplerotic netfluxes (Fig. 2). Based on these observations it was considered relevant to make simulations constraining the dehydrogenase variant to zero, in order to investigate the effect of this pathway on lysine yields. From these simulations it was seen that the maximum lysine yield dropped to 57%, when no biomass was produced (Fig. 5B). The simulation results showed that as the carbon demand for biomass production was increased, the difference in lysine yields decreased, and at Y_{sx} values higher than 40% (g biomass (g glucose)$^{-1}$) the maximum theoretical lysine yields were only marginally different, with the dehydrogenase variant being the most efficient pathway (Fig. 4A and B). At high growth rate the TCA activity required for biomass synthesis was sufficient to produce succinyl-CoA for lysine production via the succinyl branch. The succinylase variant of the lysine synthetic pathway is competing for succinylase-CoA with the TCA reaction succinyl-CoA synthease converting succinyl-CoA to succinate with the formation of an ATP. The marginally higher lysine yield for the dehydrogenase variant during high biomass formation could be explained by the missing ATP from the succinyl-CoA synthease reaction. Furthermore, it was seen that the decrease in the maximum lysine yield was due to the requirement for a higher activity of the TCA to support lysine synthesis with succinyl-CoA (Fig. 5). As a consequence more CO_2 was produced and carbon was lost. For simulations where the biomass formation was constrained to zero, the flux through the PPP was significantly lower (86%) for the ddh-negative strain (Fig. 5B), when compared to simulations where the dehydrogenase variant was used (150%) (Fig. 5A), and for the latter a cycling of the PPP was seen. The loss in carbon is due to the fact that the PPP is more efficient in NADPH generation (2 NADPH per CO_2) than the TCA cycle (1 NADPH per CO_2). If the efficiency of NADPH generation for the PPP was set to 1 mol of NADPH per CO_2, the maximum lysine yield was found to be 57% as it was the case for the ddh-negative in silico strain (data not shown), indicating that the difference seen on the maximum yield between the two strains could be explained entirely by the increased activity of the TCA cycle. For the ddh-negative in silico strain a bend of the curve was seen (Fig. 4B; gray line), as it was observed for simulations using the dehydrogenase branch, although the bend was seen earlier. The explanation for this bend was however different. In the case of the ddh-negative strain it was due to a limitation of succinyl-CoA, and not related to ATP excess as it was the case where the dehydrogenase variant was exclusively used. Based on these results, the ddh-gene looks like a promising target for overexpression by metabolic engineering. ddh has already been investigated as a potential bottleneck in lysine production. The results have been inconsistent since both positive effects (Schrumpf et al., 1991) (two to fivefold decrease in lysine accumulation when ddh was knocked out) and negative effects (Eggeling et al., 1998) (10–30% decrease by up-regulation of ddh) on lysine production by altering the flux through this pathway have been seen. For the latter no growth rate was reported for the ddh-up-regulated strain. However, for the parental strain a specific growth rate of 0.27 h^{-1} was reported. If it is assumed that the ddh-up-regulated strain had the same specific growth rate it is possible that TCA activity was sufficient to support the succinylase variant with succinyl-CoA, hence abolishing the potential positive effect of the genetic manipulation. Based on the reported simulations and the experimental work done so far on this target it would be interesting to investigate this pathway more intensively, since there is—from a theoretical point of view—a large potential for improving lysine production, at

least when we are getting closer to the theoretical maximum for this organism.

NADPH- or NADH-Dependent Glutamate Dehydrogenase

The simulation experiments comparing two different glutamate dehydrogenases were carried out according to what is described in the Simulation Methods Section, except that the NADPH-dependent glutamate dehydrogenase (*gnd*) was replaced with an NADH-dependent glutamate dehydrogenase in some cases. The objective function used was maximization for lysine production and biomass was constrained to different levels as described in the text.

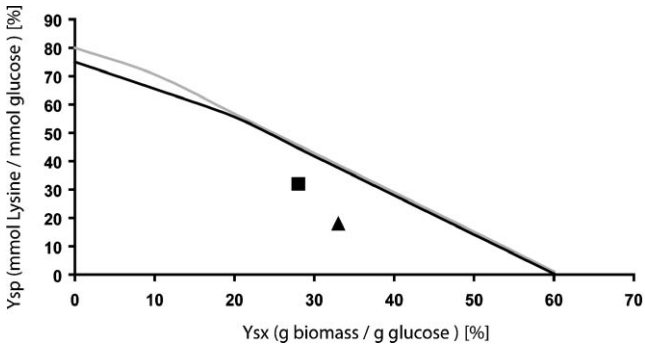

Figure 6. Simulation results maximizing for lysine production constraining biomass formation at different levels. Results shown are in silico strains (Black line: NADPH dependent glutamate dehydrogenase (*gdh*); Gray line: NADH dependent *gdh*) and experimental data from Marx et al. (1999) (square: NADPH-dependent *gdh*; Triangle: NADH-dependent *gdh*).

Marginal Improvement in Lysine Production, But Significant Redistribution of Central Carbon Metabolis, as an Effect of NADH-Dependent Glutamate Dehydrogenase

For the synthesis of one mole of lysine four molecules of NADPH is consumed. For this reason the NADPH supply in *C. glutamicum* has received a lot of attention. Marx et al. (1999) addressed this challenge by replacing the NADPH-dependent glutamate dehydrogenase (*gdh*) with an NADH-dependent *gdh*. However, this change in the metabolism of the organism did not result in improved lysine production. Instead significant redistribution of the central metabolism was observed, consequently leading to increased biomass production and lower lysine yield (Marx et al., 1999). The study made by Marx et al. (1999) was used as a case study for the genome-scale model.

Simulation experiments were carried out replacing the NADPH-dependent *gdh*-reaction with an NADH-dependent *gdh*-reaction, and the results were compared to "wild type" simulations with an NADPH-dependent *gdh*-reaction. Simulation data were further compared to experimental data.

Simulation experiment data showed general higher lysine production with the NADH-dependent *gdh*. However, the effect was only marginal when the biomass yield exceeded Y_{sx}-values of about 0.15 g biomass (g glucose)$^{-1}$ consumed (Fig. 6), which is typically the case in real fermentations. The maximum theoretical yield was increased to 80% (Fig. 6). As it was seen in the work of Marx et al. (1999), the model predicted a higher lysine yield (45%: Fig. 7A) for the in silico strain carrying the NADPH dependent *gdh*-reaction, than for the strain substituted with the NADH-dependent *gdh*-gene (39%: Fig. 7B) when the biomass formation was constrained at the same values as reported in the experiments (0.28 and 0.33 g biomass (g glucose)$^{-1}$ for the NADPH-dependent strain and NADH-dependent strain, respectively). However, the difference was lower than it was seen in the experimental work (30% for the NADPH-dependent strain and 18% for the NADH-dependent strain).

To eliminate the effect of the different biomass formation that was seen in the experimental work, simulations were carried out constraining the biomass at the same rate (0.30 g biomass (g glucose)$^{-1}$ consumed). The result gave a higher lysine yield (43%) for the NADH dependent strain when compared to the NADPH dependent strain (42%). This marginal increase was due to a small increase in Y_{sx} for the NADH-dependent strain compared to the NADPH-dependent *gdh*-strain, resulting in more carbon being available for lysine synthesis. The simulation data showed a decreased NADPH generation for the in silico strain containing the NADH-dependent *gdh* when compared to the NADPH-dependent strain (161% and 222%, respectively). As expected the moles of NADPH consumed per mole of lysine produced was lowered significantly for the NADH-dependent strain. The minimum theoretical requirement for NADPH was lowered from 4.0 to 3.0 (NADPH per lysine) in the NADH-dependent *gdh* strain (without any biomass production) when compared to the strain with an NADPH-dependent *gdh* (data not shown). Simulations using the same biomass production burdens (0.30 g biomass g glucose^{-1} consumed) also showed a lower demand for NADPH. The strain with an NADH dependent *gdh* used 3.8 mol of NADPH per mole of lysine produced, whereas the strain with the NADPH-dependent *gdh* used 5.6 mol of NADPH per mole of lysine. The experimental data from Marx et al. (1999) showed a higher NADPH demand per mole of lysine produced (7.7 mol per lysine and 6.6 mol per lysine for the NADH-dependent strain and NADPH-dependent strain, respectively). These higher values were due to the relatively low lysine yields as compared to simulations, and the lower NADPH demand for the NADPH-dependent strain can be explained by the lower flux of metabolites towards biomass formation as compared to the NADH-dependent strain.

In conclusion a significant redistribution of the central metabolism of the in silico organism is observed as an effect

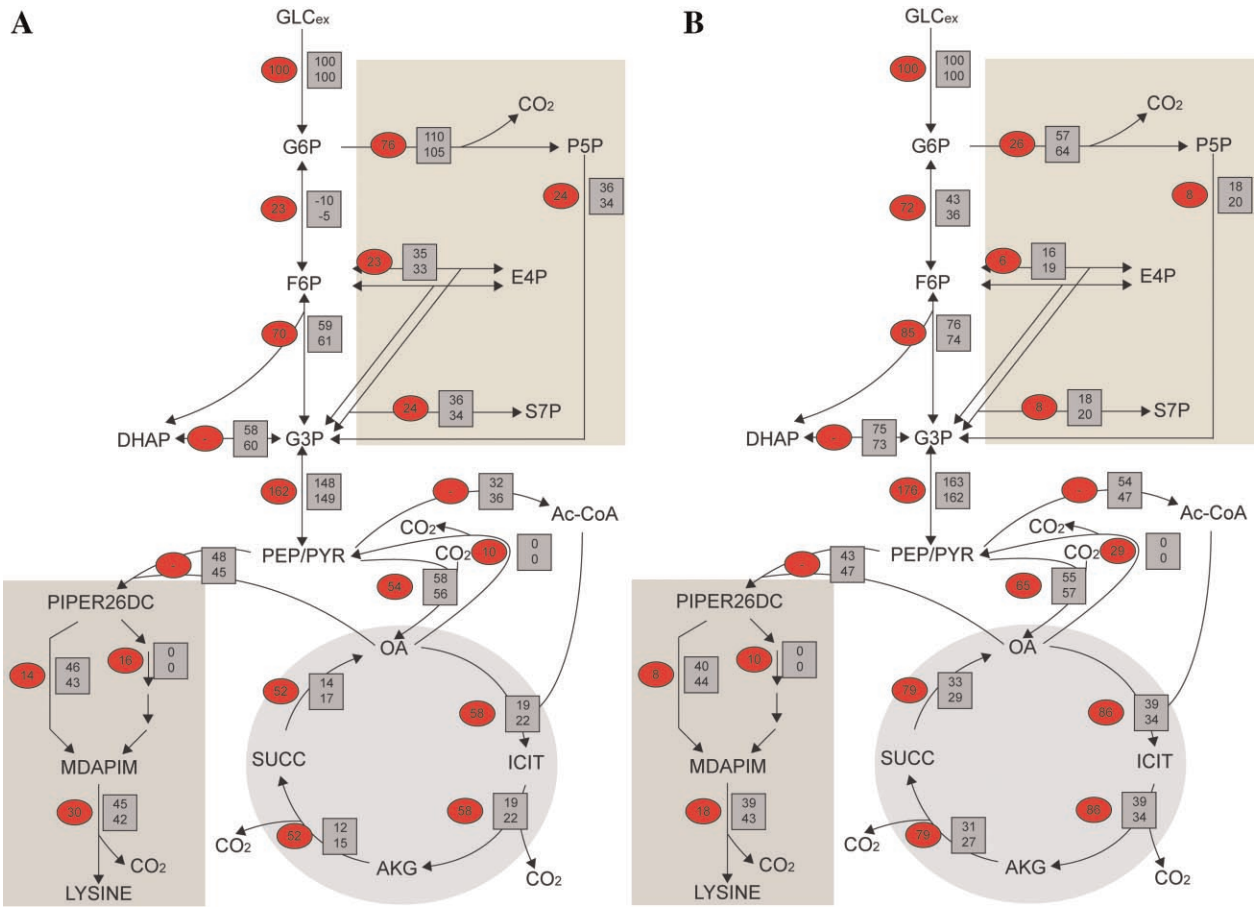

Figure 7. Flux distribution values for genome-scale model simulations (shown in gray boxes) and literature values from Marx et al. (1999) found by metabolic flux analysis (shown in red ovals). All fluxes are expressed as molar percentage of specific glucose uptake. Upper simulation value biomass production (Y_{sx}) constrained as experimental value (0.28 g biomass (glucose)$^{-1}$ for **A**) and (0.33 g biomass (glucose)$^{-1}$ for **B**). Lower simulation value Y_{sx} contained to 0.30 g biomass (glucose)$^{-1}$ for both A and B. A: Flux distribution values for experimental strain or in silico strain with NADPH dependent Glutamate dehydrogenase (GDH). B: Flux distribution values for experimental strain or in silico strain with NADH dependent GDH. Abbreviations as for Figure 1. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

of the substitution of the NADPH-dependent *gnd* with an NADH-dependent *gnd*. Although the model is unable to predict the precise change in phenotypic behavior (Increased growth and absolute change in carbon flux distributions), the simulation results are surprisingly consistent with the results of Marx et al. (1999). In addition the model predicts a higher lysine production with the NADH-dependent *gdh* strain. However, the improvement is only marginal when the biomass yield exceeds 15% of glucose consumption and an improvement on lysine yield is only seen when this is already high ($Y_{sp} > 55\%$) (Fig. 6). That NADPH is not limiting on lysine production, and that the central carbon metabolism is able to adjust to NADPH requirements of the cell is now generally accepted based on experimental studies on the regulation of the PPP (Moritz et al., 2000; Vallino and Stephanopoulos, 1994). It would however be interesting to see if this is also the case when lysine yields approaches the theoretical maximum, and some recent results have in fact indicated that this is the case, as it was seen by Becker et al. (2005) where genetic manipulations leading to an increased PPP-flux resulted in increased lysine production.

Growth of *C. glutamicum* on Lactate and Acetate

Simulations for growth on lactate and acetic acid were conducted by maximizing for growth as the objective function. Simulations were carried out according to what is described in the Simulation Methods Section, except no constraints were made on the two reactions catalyzed by malic enzyme and oxaloacetate decarboxylase. In some cases constraints were made on individual reactions, which are described in the text.

C. glutamicum is able to grow on various substrates including a number of organic acids such as lactate (Cocaign-Bousquet and Lindley, 1995) and acetate (Wendisch et al., 2000). Simulations were carried out using

acetate and lactate as carbon source, and the results were compared to results found in literature.

Lactate as Carbon Source

Simulations maximizing for growth was carried out when lactate was used as a carbon source (Fig. 8). As expected the in silico organism could utilize lactate as a carbon source. The maximum biomass yield on substrate was 0.54 g biomass (g lactate)$^{-1}$, which was lower than when glucose was used as a substrate (0.61 g biomass (g glucose)$^{-1}$). The difference was due to the lower efficiency in ATP- and NADPH synthesis, when the gluconogenetic pathways were included in the model. Experimental data for a carbon limited chemostat showed a biomass yield of 0.36 g biomass (g lactate)$^{-1}$ ($D = 0.17$ h^{-1}) whereas the yield increased as dilution rate was increased (Cocaign-Bousquet and Lindley, 1995). It needs to be emphasized that in the biomass yield calculations made by Cocaign-Bousquet and Lindley (1995) the pyruvate produced as a byproduct was subtracted from the substrate carbon, so the biomass yield on lactate including pyruvate production in reality was lower. For a dilution rate of 0.28 h^{-1} the biomass yield was 0.42 g biomass (g lactate)$^{-1}$ (0.3 g biomass (g lactate)$^{-1}$ when pyruvate production was taken into account). At this dilution rate the chemostat was not carbon limited, and an overflow of pyruvate was seen resulting in excretion of this compound. At the organisms maximum growth rate ($\mu = 0.35$ h^{-1}) a biomass yield of 0.61 g biomass (g lactate)$^{-1}$ (0.4 g biomass (g lactate)$^{-1}$ when pyruvate production was included) was seen, and the efflux of pyruvate was further increased. When comparing experimental data with simulation data, it was seen that the model data (Fig. 8; gray boxes) was comparable to the data for the carbon-limited chemostat with $\mu = 0.17$ h^{-1} (Fig. 8; red ovals 1st flux from the top), whereas for data for the maximum growth rate ($\mu = 0.35$ h^{-1}), fluxes around phosphoenolpyruvate, pyruvate, oxaloacetate and malate were significantly different (Fig. 8; red ovals 2nd flux from top). In addition to this an efflux of pyruvate (21% of lactate uptake), and a reduced flux through the TCA was observed (Cocaign-Bousquet and Lindley, 1995). Looking more into details of the experimental data at high growth rates, and hence high substrate uptake rates, growth rate limitations within the central metabolism were observed (Cocaign-Bousquet and Lindley, 1995). During these conditions some enzymes became rate limiting, and the organism found alternative pathways which could not be predicted by the model.

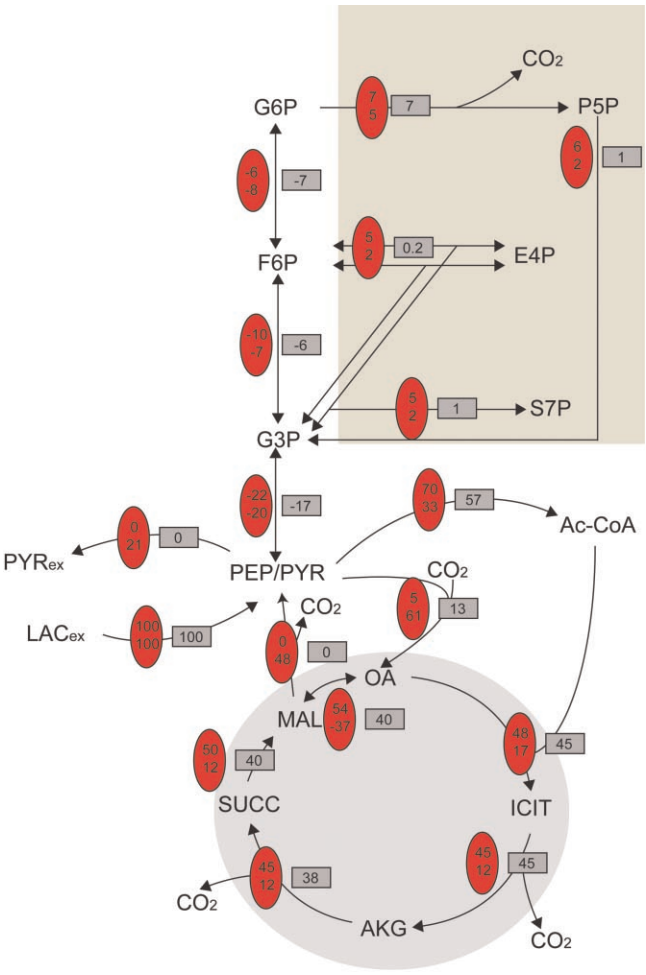

Figure 8. Metabolic flux distribution of *C. glutamicum* growing on lactate and maximizing for growth. All fluxes are expressed as molar percentage of the specific uptake rate of lactate. Numbers in red ovals are taken from Cocaign-Bousquet and Lindley (1995). Upper flux $\mu = 0.17$ h^{-1}; lower flux $\mu = 0.35$ h^{-1}. Numbers in gray boxes are results from simulations with in silico organism. LAC_{ex}: extracellular lactate; PYR_{ex}: extracellular pyruvate. Other abbreviations see Figure 1. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

Maximization for growth was simulated during growth on either acetate as single carbon source (Fig. 9), or during

Acetate as Carbon Source

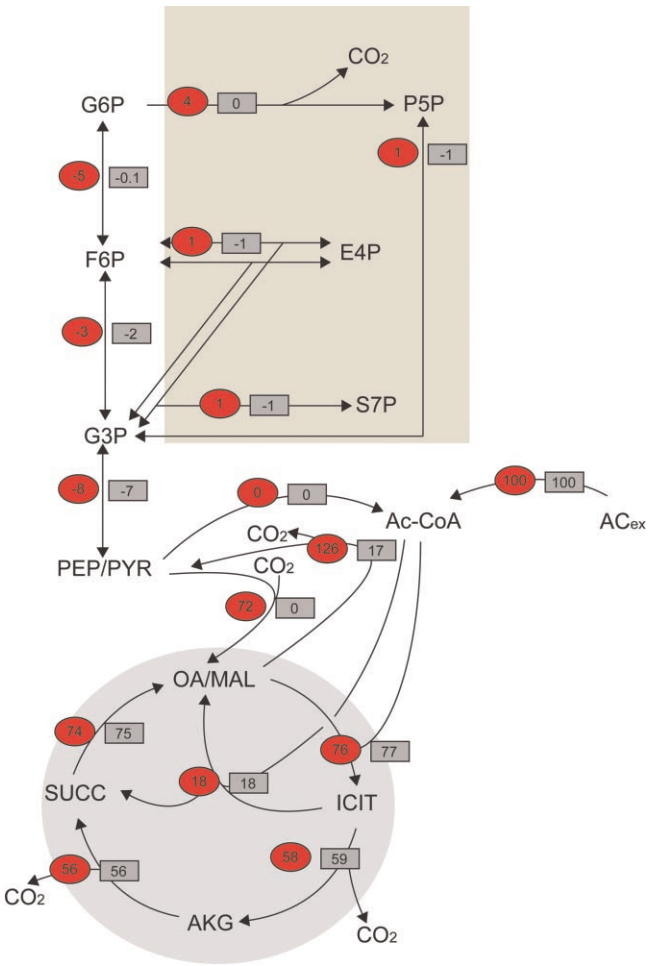

Figure 9. Metabolic flux distribution of *C. glutamicum* growing on acetate and maximizing for growth. All fluxes are expressed as molar percentage of the specific uptake rate of acetate. Numbers in red ovals are taken from Wendisch et al. (2000). Numbers in gray boxes are results from model simulations. AC_{ex}: extracellular acetate. Other abbreviations see Figure 1. [Color figure can be seen in the online version of this article, available at www.interscience.wiley.com.]

simultaneous acetate and glucose consumption (data not shown). The maximum biomass yield on acetate (0.48 g biomass (g acetate)$^{-1}$) was found to be lower than on glucose (0.61 g biomass (g glucose)$^{-1}$), and the glyoxylate shunt was observed to be active. Both observations were consistent with experimental data (Wendisch et al., 2000), although biomass yields for the experimental data were lower being 0.24 g biomass (g acetate)$^{-1}$ and 0.34 g biomass (g glucose)$^{-1}$ for acetate and glucose, respectively.

It was observed that no growth was possible for the in silico organism when the glyoxylate shunt was shut down (data not shown), which was also observed by Wendisch et al. (2000). During optimal growth the model predicted a flux through the glyoxylate cycle corresponding to 25% of the substrate uptake (Fig. 9). This flux was 18% for the experimental data of Wendisch et al. (2000). Further investigation of the glyoxylate shunt branch point showed, that when the carbon flux through the glyoxylate shunt was constrained to 18% of the acetate uptake, the biomass yield was lowered to 0.35 g biomass (g acetate)$^{-1}$ (data not shown). When the carbon flux through the glyoxylate shunt was increased by constraining this flux, a decrease in the biomass yield was seen (data not shown), indicating that the carbon flux predicted by the model is truly a maximum for the in silico organism.

The model predicted that the flux through the PPP was entirely omitted, and NADPH needed for biomass was solely generated by the isocitrate dehydrogenase reaction, whereas the data of Wendisch et al. (2000) showed a low flux (4%) through the PPP. When PPP-flux was constrained to 4% simultaneously with the glyoxylate shunt constrain above, the biomass yield further decreased to 0.33 g biomass (g acetate)$^{-1}$. The TCA-flux was higher for the experimental data as it was the case for the model.

Simulations for co-metabolism of acetate and glucose were carried out (data not shown). Wendisch et al. (2000) found that the glyoxylate cycle was active during co-metabolism of these two substrates, which was a surprise since this was not seen for *E. coli* cells grown under the same conditions (Walsh and Koshland, 1985), and since this pathway is not essential when glucose is present. The model predicted the same behavior when substrate uptake was constrained at the same levels (data not shown). However, when the glyoxylate cycle was removed from the model the growth yield of the organism was not altered (data not shown), which in practice means that from a stoichiometric point of view it is insignificant which route is used under the given conditions. The preferred use of the glyoxylate route in Wendisch et al. (2000) is probably due to acetate specific induction, which earlier has been seen for *C. glutamicum* during these conditions (Gerstmeier et al., 2003). Anaplerotic fluxes were higher for the experimental data than for the model, and a significant cycling between C3 and C4 was seen (Fig. 9) resulting in an anaplerotic netflux of 54% towards the pyruvate/phosphoenolpyruvate pool which was higher than for the model (17%) (Fig. 9).

Conclusions

A validated metabolic network of *C. glutamicum* ATCC 13032 was constructed, and by using flux-balance analysis the in silico model was able to predict metabolic fluxes during lysine production and growth under various conditions, which were consistent with experimental values. This work also showed that genome scale metabolic model simulations needs to be combined with data from other sources, that is, flux data or transciptomic data, in order to improve the prediction power of the model. The model was able to predict potential targets for metabolic engineering for improving lysine production in *C. glutamicum*, and hence serves as a useful tool for future directing of metabolic

engineering strategies resulting in improved lysine production. The model also serves as an extensive compendium on *C. glutamicum* metabolism, and it is our hope that by combining the model with datasets from high throughput experimental techniques such as transcriptomics, fluxomics and metabolomics, the prediction power of the model can be further improved.

Ana Paula Oliveira is acknowledged for help and guidance during model construction and for introducing the Biooptv4.9 software.

References

Åkesson M, Förster J, Nielsen J. 2004. Integration of gene expression data into genome-scale metabolic models. Metab Eng 6:285–293.

Becker J, Klopprogge C, Zelder O, Heinzle E, Wittmann C. 2005. Amplified expression of fructose 1,6-bisphosphatase in *Corynebacterium glutamicum* increases *in vivo* flux through the pentose phosphate pathway and lysine production on different carbon sources. Appl Environ Microbiol 71:8587–8596.

Bonarius HPJ, Schmid G, Tramper J. 1997. Flux analysis of underdetermined metabolic networks: The quest for the missing constraints. Trends Biotechnol 15:308–314.

Borodina I, Krabben P, Nielsen J. 2005. Genome-scale analysis of *Streptomyces coelicolor* A3(2) metabolism. Genome Res 15:820–829.

Bott M, Niebisch A. 2003. The respiratory chain of *Corynebacterium glutamicum*. J Biotechnol 104:129–153.

Brand S, Niehaus K, Pühler A, Kalinowski J. 2003. Identification and functional analysis of six mycolyltransferase genes of *Corynebacterium glutamicum* ATCC 13032: The genes *cop1*, *cmt1*, and *cmt2* can replace each other in the synthesis of trehalose dicorynomycolate, a component of the mycolic acid layer of the cell envelope. Arch Microbiol 180:33–44.

Cocaign-Bousquet M, Lindley ND. 1995. Pyruvate overflow and carbon flux within the central metabolic pathways of *Corynebacterium glutamicum* during growth on lactate. Enzyme Microb Technol 17:260–267.

Cocaign-Bousquet M, Guyonvarch A, Lindley ND. 1996. Growth rate-dependent modulation of carbon flux through central metabolism and the kinetic consequences for glucose-limited chemostat cultures of *Corynebacterium glutamicum*. Appl Environ Microbiol 62:429–436.

Collins MD, Doodfellow M, Minnikin DE. 1982a. A survey of the structures of mycolic acids in *Corynebacterium glutamicum* and related taxa. J Gen Microbiol 128:129–149.

Collins MD, Goodfellow M, Minnikin DE. 1982b. Fatty acid composition of some mycolic acid-containing coryneform bacteria. J Gen Microbiol 128:2503–2509.

Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BØ. 2004. Integrated high-throughput and computational data elucidates bacterial networks. Nature 429:92–96.

Cremer J, Eggeling L, Sahm H. 1991. Control of the lysine biosynthesis sequence in *Corynebacterium glutamicum* as analysed by overexpression of the individual corresponding genes. Appl Environ Microbiol 57: 1746–1752.

Daffé M. 2005. The cell envelope of corynebacteria. In: Eggeling L, Bott M, editors. Handbook of *Corynebacterium glutamicum.* Boca Raton: CRC Press. p 121–148.

de Graaf AA. 2000. Metabolic flux analysis of *Corynebacterium glutamicum*. In: Schügerl K, Bellgard KH, editors. Bioreaction engineering: Modelling and control. Berlin: Springer Verlag. p 506–555.

Edwards JS, Palsson BØ. 2000. The *Escherichia coli* MG1655 *in silico* metabolic genotype: Its definition, characteristics, and capabilities. Proc Natl Acad Sci 97:5528–5533.

Edwards JS, Ibarra RU, Palsson BØ. 2001. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nat Biotechnol 19:125–130.

Eggeling L, Oberle S, Sahm H. 1998. Improved L-lysine yield with *Corynebacterium glutamicum*: Use of *dapA* resulting in increased flux combined with growth limitation. Appl Microbiol Biotechnol 49:24–30.

Feist AM, Scholten JCM, Palsson BØ, Brockman FJ, Ideker T. 2006. Modelling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. Mol Syst Biol 2:1–14.

Förster J, Famili I, Fu P, Palsson BØ, Nielsen J. 2003. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. Genome Res 13:244–253.

Gerstmeier R, Wendisch VF, Schnicke S, Ruan H, Farwic M, Reinscheid D, Eikmanns BJ. 2003. Acetate metabolism and the regulation in *Corynebacterium glutamicum*. J Biotechnol 104:99–122.

Heinemann M, Kümmel A, Ruinatscha R, Panke S. 2005. In silico genome-scale reconstruction and validation of the Staphylococcus aureus metabolic network. Biotechnol Bioeng 92:850–864.

Herrgård MJ, Fong SS, Palsson BØ. 2006. Identification of Genome-scale metabolic network models using experimentally measured flux profiles. PLoS Comput Biol 2:676–686.

Hoischen C, Krämer R. 1990. Membrane alteration is necessary but not sufficient for effective glutamate secretion in *Corynebacterium glutamicum*. J Bacteriol 172:3409–3416.

Ingraham JL, Maaløe O, Niedhardt FC. 1983. Growth of the bacterial cell. Sunderland, Massachusetts: Sinauer Associates, Inc. p 435.

Jang KH, Pierotti D, Kemp GW, Best GR, Britz ML. 1997. Mycolic acid composition of *Corynebacterium glutamicum* and its cell surface mutants: Effects of growth with gycine and isonicotinic acid hydrazide. Microbiology (UK) 143:3209–3221.

Kalinowski J, Bathe B, Bartels D, Bischoff N, Bott M, Burkovski A, Dusch N, Eggeling L, Eikmanns BJ, Gaigalat L. 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. J Biotechnol 104:5–25.

Keddie RM, Cure GL. 1978. Cell-wall composition of coryneform bacteria. In: Bousfield IJ, Calley AG, editors. Coryneform bacteria. London: Academic Press. p 47-84.

Kelle R, Hermann T, Bathe B. 2005. L-Lysine production. In: Eggeling L, Bott M, editors. Handbook of *Corynebacterium glutamicum*. Boca Raton: CRC Press. p 465–488.

Krömer JO, Sorgenfrei O, Klopprogge K, Heinzle E, Wittmann C. 2004. In-depth profiling of lysine-producing *Corynebacterium glutamicum* by combined analysis of the transcriptome, metabolome, and fluxome. J Bacteriol 186:1769–1784.

Marx A, de Graaf AA, Wiechert W, Eggeling L, Sahm H. 1996. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. Biotechnol Bioeng 49:111–129.

Marx A, Striegel K, de Graaf AA, Sahm H, Eggeling L. 1997. Response of the central metabolism of *Corynebacterium glutamicum* to different flux burdens. Biotechnol Bioeng 56:168–180.

Marx A, Eikmanns BJ, Sahm H, de Graaf AA, Eggeling L. 1999. Response of the central metabolism in *Corynebacterium glutamicum* to the use of an NADH-dependent glutamate dehydrogenase. Metab Eng 1:35–48.

Moritz B, Striegel K, de Graff AA, Sahm H. 2000. Kinetic properties of the glucose-6-phosphate and 6-phosphogluconate dehydrogenases from *Corynebacterium glutamicum* and their application for predicting pentose phosphate pathway flux in vivo. Eur J Biochem 267:3442–3452.

Oliveira AP, Nielsen J, Förster J. 2005. Modelling *Lactococcus lactis* using a genome-scale flux model. BMC Microbiol 5:39.

Palsson BØ. 2000. The challenges of *in silico* biology. Nat Biotechnol 18: 1147–1150.

Palsson BØ. 2006. Systems biology: Properties of reconstructed networks. New York: Cambridge University Press. p 322.

Patil KR, Åkesson M, Nielsen J. 2004. Use of genome-scale microbial models for metabolic engineering. Curr Opin Biotechnol 15:64–69.

Petersen S, de Graff AA, Eggeling L, Mollney M, Wiechert W, Sahn H. 2000. In vivo quantification of parallel and bidirectional fluxes in the anaplerosis of *Corynebacterium glutamicum*. J Biol Chem 275:35932–35941.

Price ND, Papin JA, Schilling CH, Palsson BØ. 2003. Genome-scale microbial *in silico* models: The constrain-based approach. Trends Biotechnol 21:162–169.

Puech V, Chami M, Lemassu A, Laneelle MA, Schiffler B, Gounon P, Bayan N, Benz R, Daffé M. 2001. Structure of the cell envelope of corynebacteria: Importance of the non-covalently bound lipids in the formation of the cell wall permeability barrier and fracture plane. Microbiology 147:1365–1382.

Reed JR, Vo TD, Schilling CH, Palsson BØ. 2003. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). Genome Res 4:R54.

Schilling CH, Palsson BØ. 2000. Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. J Theoret Biol 203:249–283.

Schilling CH, Schuster S, Palsson BØ, Heinrich R. 1999. Metabolic pathway analysis: Basic concepts and scientific applications in the post-genomic era. Biotechnol Prog 15:296–303.

Schilling CH, Covert MW, Famili I, Church GM, Edwards JS, Palsson BØ. 2002. Genome-scale metabolic model of *Helicobacter pylori* 26695. J Bacteriol 184:4582–4593.

Schrumpf B, Schwarzer A, Kalinowski J, Pühler A, Eggeling L, Sahn H. 1991. A functionally split pathway for lysine synthesis in *Corynebacterium glutamicum*. J Bacteriol 173:4510–4516.

Sonntag K, Eggeling L, de Graaf AA, Sahm H. 1993. Flux partitioning in the split pathway of lysine synthesis in *Corynebacterium glutamicum*. Quantification by 13C- and 1H-NMR spectroscopy. Eur J Biochem 213:1325–1331.

Sonntag K, Schwinde JW, de Graaf AA, Marx A, Eikmanns BJ, Wiechert W, Sahm H. 1995. 13C NMR studies of the fluxes in the central metabolism of *Corynebacterium glutamicum* during growth and overproduction of amino acids in batch cultures. Appl Microbiol Biotechnol 44:489–495.

Stephanopoulos G, Vallino JJ. 1991. Network rigidity and metabolic engineering in metabolite overproduction. Science 252:1675–1681.

Stephanopoulos G, Aristidou AA, Nielsen J. 1998. Metabolic engineering—Principles and methodologies. California, San Diego: Academic Press. p 725.

Strelkov S, Elstermann M, Schomburg D. 2004. Comprehensive analysis of metabolites in *Corynebacterium glutamicum* by gas chromatography/mass spectromety. J Biol Chem 385:853–861.

Teusink B, Wiersma A, Molenaar D, Francke C, de Vos WM, Siezen RJ, Smid EJ. 2006. Analysis of growth of *Lactobacillus plantarum* WCFS1 on a complex medium using a genome-scale model. J Biol Chem 281:40041–40048.

Vallino JJ, Stephanopoulos G. 1993. Metabolic flux distribution in *Corynebacterium glutamicum* during growth and lysine overproduction. Biotechnol Bioeng 41:633–646.

Vallino JJ, Stephanopoulos G. 1994. Carbon flux distributions at the glucose 6-phosphate branch point in *Corynebacterium glutamicum* during lysine overproduction. Biotechnol Prog 10:327–334.

Varela CA, Baez ME, Agosin E. 2004. Osmotic stress response: Quantification of cell maintenance and metabolic fluxes in a lysine-overproducing strain of *Corynebacterium glutamicum*. Appl Environ Microbiol 70: 4222–4229.

Walsh K, Koshland DE. 1985. Branch point control by the phosphorylation state of isocitrate dehydrogenase. J Biol Chem 260:8430–8437.

Wendisch VF, de Graaf AA, Sahm H, Eikmanns BJ. 2000. Quantitative determination of metabolic fluxes during co-utilization of two carbon sources: Comparative analyses with *Corynebacterium glutamicum* during growth on acetate and/or glucose. J Bacteriol 182:3088–3096.

Wittmann C, Heinzle E. 2001. Application of MALDI-TOF MS to lysine-producing *Corynebacterium glutamicum*: A novel approach for metabolic flux analysis. Eur J Biochem 268:2441–2455.

Wittmann C, Heinzle E. 2002. Genealogy profiling through strain improvement by using metabolic network analysis: Metabolic flux genealogy of several generations of lysine-producing *Corynebacteria*. Appl Environ Microbiol 68:5843–5859.