Quantitative Physiology of *S. cerevisiae* using Metabolic Network Analysis

Hans Christian Lange
1. The possibility of a shuttle exporting NADH out of the mitochondria and using cytosolic and mitochondrial malate dehydrogenases can not be excluded by the fact that the mitochondrial reaction proceed from malate towards oxaloacetate.

2. The existence of three genes in S. cerevisiae encoding for isoenzymes of malate dehydrogenase in the cytosol, the mitochondria, and the peroxisome makes this enzyme an unsuitable choice as marker for intracellular localization studies in combination with enzyme activity assays.

3. The dependence on the growth substrate of the intracellular localization of fumarase is a clear, but often ignored sign, that the growth conditions are essential during localization studies of enzymes and should be taken into account when interpreting the results.

4. The summing up to 100% is not a guarantee, that no item was missed during the procedure.

5. The conclusion “A null Cit2 mutant is perfectly healthy” based on few growth experiments reveals a simplistic attitude toward the complexity of life or a strange concept of health.

6. The use of a hsp2Δ strain in the research on an oxaadipate carrier points to either an accidental finding or a bad planning of the experiment.

7. The claim of a peroxisomal localization for the citrate synthase encoded by the CIT2 gene of S. cerevisiae shows that researchers tend to find what they are looking for.

8. The improvement of crop yields by what ever means will not eradicate hunger.

9. Heisenberg’s uncertainty principle is despite its popularity the most undervalued discovery in physics of the last century

10. Freedom and safety do not go together.
Stellingen

behorende bij het proefschrift
“Quantitative Physiology of *Saccharomyces cerevisiae* Based on Metabolic Network Analysis”
door Hans Christian Lange

1. De mogelijkheid van een shuttle die NADH uit de mitochondria exporteert met behulp van cytosolaire en mitochondriële malaat dehydrogenasen, kan niet worden uitgesloten door het feit dat de mitochondriële reactie van malaat richting oxaloacetaat verloopt.

2. Het bestaan van drie genen in *S. cerevisiae*, die voor verschillende isoenzymen van malaat dehydrogenase in de cytosol, de mitochondria en in het peroxisoom coderen, maken dit enzym ongeschikt als merker voor het bestuderen van de intracellulaire localisatie in combinatie met een enzymactiviteitstest.

3. De afhankelijkheid van de intracellulaire localisatie van fumarase van het mediumsubstraat is een duidelijk, maar vaak genegeerd, teken dat de groeiomstandigheden van belang zijn, wanneer de localisatie van enzymen wordt onderzocht. Hier moet rekening mee worden gehouden bij de interpretatie van de resultaten.

4. Wanneer een opsomming een totaal van 100% oplevert, is dat nog geen garantie dat er niets is gemist.

5. De op slechts enkele groei experimenten gebaseerde conclusie "een null Cit2 mutant is volledig gezond" geeft blijk van een heel eenvoudige kijk op de complexiteit van het leven of een merkwaardig idee van gezondheid.

6. Het gebruik van een *fus2Δ* strain in het onderzoek naar een oxoadipaat transporter wijst op een toevallige vinding of een slechte planning van het experiment.

7. De claim van een peroxisomale locatie van de citraat synthase gecodeerd door het CIT2 gen van *S. cerevisiae* toont aan dat onderzoekers plegen te vinden waarnaar ze zoeken.

8. De opbrengstverhoging van planten lost het wereldvoedselprobleem niet op.

9. Ondanks zijn populariteit, is Heisenbergs onzekerheidsprincipe de meest ondergewaardeerde ontdekking van de afgelopen eeuw.

10. Vrijheid en veiligheid gaan niet samen.
QUANTITATIVE PHYSIOLOGY OF
SACCHAROMYCES CEREVISIAE
BASED ON
METABOLIC NETWORK ANALYSIS

HANS CHRISTIAN LANGE
QUANTITATIVE PHYSIOLOGY
OF SACCHAROMYCES CEREVISIAE
BASED ON
METABOLIC FLUX ANALYSIS

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op maandag 11 maart 2002 om 13:30 uur

door

HANS CHRISTIAN LANGE

Diplom-Ingenieur der Verfahrenstechnik,
Technische Universität Hamburg-Harburg

geboren te Bad Oldesloe, Duitsland
Dit proefschrift is goedgekeurd door de promoter:

Prof. dr. ir. J. J. Heijnen

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof. dr. ir. J. J. Heijnen
Prof. Dr. techn. J. Nielsen MSc
Prof. Dr.-Ing. M. Reuss
Prof. dr. ir. C.T. Verrips
Prof. dr. ir. G.W.K. van Dedem
Prof. dr. J. P. van Dijken
Prof. dr. J. T. Pronk

Technische Universiteit Delft, promoter
Danmarks Tekniske Universitet
Universität Stuttgart
Universiteit Utrecht
Technische Universiteit Delft
Technische Universiteit Delft
Technische Universiteit Delft

The studies presented in this thesis were performed at the Kluyver Laboratory of Biotechnology, Delft University of Technology. The research was financially supported by the European Commission (Framework IV: “From gene to product in yeast: a quantitative approach”).

# Table of content

**chapter 1** General Introduction ................................................. 1

**chapter 2** Statistical Reconciliation of the Biomass Composition .......... 19

**chapter 3** Characterisation of null mutants of glyoxylate cycle and  
gluconeogenic enzymes ................................................................. 43

**chapter 4** Setup of a compartmented metabolic network model .............. 65

**chapter 5** Energetic aspects of the compartmentation of the metabolism ... 99

**chapter 6** Metabolic flux analysis of a *tpiΔ* deletion mutant ............... 125

**chapter 7** Rapid sampling for intracellular metabolites determination ...... 153

**chapter 8** Summary - Samenvatting .............................................. 175
Chapter 1

General Introduction
Introduction

Developments in life sciences allows nowadays the directed manipulation of the cell blueprint, the DNA. Molecular biologists are capable of elimination or addition of single and multiple genes in a controlled fashion; controlled expression of specific genes at a desired moment lies in the realm of feasibility. The number of completely sequenced genomes rises weekly, while the required time dropped from years to mere month, improvements in the sequencing capacities comparable to the ones seen in the development of computer chips are likely. Humanity will soon be able to unravel the genetic code of any given organism in a matter of hours or days, creating vast amounts of data waiting to be converted into information.

The binary information contained in the DNA sequence is successfully used especially in the pharmaceutical industry for the production of complex molecules, such as antibiotics and hormones. Various strategies are followed to change this information and obtain a reprogrammed microorganism with more desirable features. Microbial pathways leading to desired compounds are over-expressed or transferred into another organism better suited to the conditions of large scale fermentations. Pathways leading to unwanted byproducts are eliminated to force a larger share of the used starting material into the desired product. However, yields of these processes are rarely above few milli-mol per mol of substrate, mostly high energy containing compounds such as sugars or relatively expensive amino acids. Low yields are economically viable for high value added compounds used in the fine chemistry and pharmaceutical industry, but even here limitation in resources and competition will ultimately force the optimization of these processes.

The low yields of the molecules are the result of complex formation pathways, which evolved in nature over a time span of million years under the constrains to optimize the survival and replication of the host microorganism. This has imminent implications for the structure of the metabolic network, which to elucidate and understand will be the task for the near future. Only with this knowledge, we will be able to further optimize the production processes, where the cell’s objective should be reversed: the formation of a specific compound is the main goal, biomass becomes an undesired byproduct.

It requires the cooperation of very different scientific fields to overcome these problems and to harness the potentials of microbiology for the benefit for humanity. While the understanding of the network’s functions falls in the realm of the microbiologist, the analysis of the network behaviour is a task more suitable for a mathematician, the engineer is required, where a quantitative description is the desired outcome, as it is the case in the industrial development of microbiological derived products. This cooperation has led to the emergence of the field of metabolic engineering (Bailey, 1991), which focuses on the development of the mathematical tools required for a successful reprogramming of microorganisms.
Mathematical tools for the determination of metabolism

The development of the biological tools for the manipulation of the DNA quickly lead to the question, which of the genes have to be manipulated to obtain the desired traits in the modified organism. To address this question, new mathematical tools needed to be developed (Niederberger et al., 1992). The complexity of the mathematical problem is a result of the cell’s dynamic structure: “any perturbation of the cell will result in a multigene/multitranscript/multiprotein response” (Bailey, 1999); known for a long time, but by no means common knowledge yet, the understanding of the cell’s metabolism requires thus a global view on the complex interactions of the involved components, and can not be achieved simply in terms of a few steps taken out of their metabolic context (Bowden, 1999). The future lies, therefore, in complex, holistic mathematical models of the metabolism, which will include knowledge from all different organisational levels, from the genome down to the metabolome (Bailey, 1998). Tools to accomplish this tasks will be a combination of physiological studies, metabolic flux analysis, metabolic control analysis, thermodynamic descriptions of the pathways, and kinetic models (Nielsen, 1998).

\[
\begin{bmatrix}
1 & 1 & 0 \\
1 & 0 & 0 \\
0 & -1 & 1 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
r_1 \\
r_2 \\
r_3 \\
r_4
\end{bmatrix}
= 
\begin{bmatrix}
r_A \\
r_B \\
r_D
\end{bmatrix}
\]

Figure 1: metabolic network with its mathematical representation; all internal fluxes \( r_i \) can be determined for (A), if two of the exchange fluxes \( r_A, r_B, \) and \( r_D \) are known; example (B) does not allow the independent determination of the internal fluxes \( r_i \) if only \( r_A \) and \( r_D \) are known.

Already in the earlier 1980s, the concept of the metabolic flux analysis was developed. Knowledge of the metabolic pathways, describing at least the essential central reactions of the cell, was combined with measurements of uptake rates. The balancing of metabolites (figure 1) was then used to determine the fluxes through all reactions of the cell’s metabolic network (Bonnet et al., 1980; Holms, 1986; Holms, 1996). Early successful applications were the overproduction of amino acids, e.g. the lysine production in
Corynebacterium glutaminum (Vallino and Stephanopoulos, 1993). Mathematical tools for the analysis of the applied metabolic networks were developed to test for the identifiability of fluxes, to search for conserved moieties and parallel pathways, and to evaluate the sensitivity of the network (Schilling and Palsson, 1998; van der Heijden et al., 1994). The development of the elemental mode analysis enabled the detection of all possible product routes from the total of metabolic reactions (Schuster et al., 1999). Metabolic flux analysis was consequently used successfully to e.g. quantify the maximal theoretical yields (van Gulik and Heijnen, 1995) and estimate the energetics of the cell's metabolism (Vanrolleghem et al., 1996; Verduyn, 1992).

However, several questions around the metabolic flux analysis remained unsolved, one of the most significant one being the identification of the fluxes through parallel routes. The maximum yield principle, which assumes a metabolic network structure maximizing the biomass yield on the given substrate, proved to be a valuable tool to choose from several network option, and was successfully applied in several works (van Gulik and Heijnen, 1995; Vanrolleghem et al., 1996; Varner and Palsson, 1994) and is also being used frequently in this study. The quantification of parallel pathways was partly made possible by the introduction of the co-factor balance. However, the aerobic cultivation of an organism requires a thorough, quantitative understanding of the energy metabolism around the respiratory chain to allow the use of the co-factor balances for ATP. Aerobic cultivation further leads to the biomass being the major product of the metabolic network, thus its composition becomes a major factor in the determination of metabolic fluxes, a fact often overlooked and neglected.

The metabolic flux analysis was further improved and broadened in it reach through the inclusion of the isotope balances (figure 2), which allowed a more precise determination of parallel and bidirectional fluxes, which were unaccessible to the metabolite based flux analysis (Schmidt et al., 1997). Based on the isotope distribution in the metabolite pools analysis of nuclear magnetic resonance (NMR) spectra, this tool also allowed the identification of the topology of the network, and should thus be termed more appropriately metabolic network analysis (Christensen and Nielsen, 1999).

Parallel to the development of the metabolic flux calculations for stoichiometric analysis, the metabolic control analysis (MCA) emerged as a new field for kinetic analysis (Fell, 1992; Heinrich and Rapoport, 1974; Small and Kacser, 1993a; Small and Kacser, 1993b), when it became obvious, that the simple over-expression of a single enzyme in the metabolic pathway would not result in the hoped for improvements (Fell, 1998). Here, an attempt was made to formalise the influence of a single enzyme on the fluxes throughout
the network. The obtained control coefficient describes the response of the metabolic fluxes to small changes in the enzymatic activity. MCA was able to theoretically explain the painfully often observed absence of the "bottleneck" enzyme and the meagre improvements of overexpression experiments. However, in cells with a genetically modified pathway, the enzyme activities usually change within orders of magnitude, thus severely limiting the applicability of the MCA. Attempts to overcome this limitation relies on the simplified presentation of the kinetics of the underlying network (Small and Kacser, 1993a; Small and Kacser, 1993b). Recent studies tried to link the MCA with the flux analysis to improve it applicability (Heijnen, 2000; Liao and Delgado, 1998).

Although, the development of the mathematical tools was accompanied by an increasing knowledge of the underlying biology, many uncertainties remain. The availability of the complete genome for many organisms theoretically allows the complete description of the metabolic activities, however, the function of most open reading frames (ORFs) has yet to be identified. Even for a well studied organism, such as baker’s yeast, only about 60% of the proteins have been assigned a specific function. The increasing amount of data allowed the inclusion of a mitochondrial compartment in the latest metabolic models (Granström et al., 2000; Nissen et al., 1997; Ostergaard et al., 2001; Wahlbom et al., 2001). However, the precise distribution of the complete metabolic network over the various compartments together with the appropriate intracellular transport systems had not been addressed in a systematic way. Furthermore, many questions about the exact structure of the metabolic network remain unanswered. Here, metabolic flux analysis with different conceivable networks can help to identify the most likely network options, if combined with proper measurement data from closely related cultivations. In this way, the mathematical analysis contributes to the wealth of biological information obtain with more traditional biochemical tools.

Currently, the focus is shifting from the stoichiometry and the modelling of steady state fluxes towards the kinetics and dynamic modelling of the networks. Such a model would allow, for example, the mathematical determination of the control coefficients under various conditions, and thus being of high value for identification of targets for genetic modifications. Successful attempts were made for single pathways, e.g. the glycolysis in yeast (Rizzi et al., 1997; Visser et al., 2000) and ethanol production in E. coli (Hatzimanikatis et al., 1997b). Two main strategies are followed, which differ in their requirements for knowledge about the enzyme kinetics. While Rizzi et al. used the complete set of kinetic data on each reaction step, the groups around Visser and Hatzimanikatis both minimized the number of the model’s parameters using an approximative kinetics for of each reaction. The latter strategy is thus applicable for pathways, which are sparsely characterized and with that allows faster setup of the mathematical model, but, at the same time, has a more limited predictive range.

Verication of any mathematical model will require reliable data. Despite the large progress, the precise quantification of metabolic processes remains a challenging task, be it the measurement of gene expression levels, the determination of catalytic fluxes through enzymes, or the measurement of intracellular concentrations (Rizzi et al., 1997; Schäfer et al., 1999). To facilitate the next step towards a dynamic model of the metabolism, reliable methods for the precise quantification of intracellular metabolite levels are a prerequisite.
The deepening of our understanding of metabolic processes and the desired predictive modelling of these processes requires the constant improvement of the mathematical tools through the inclusion of new microbiological knowledge into these models, the broadening of their predictive capacities, and the feedback of thus obtained insight towards the knowledge on microbial metabolism. This thesis deals with the qualitative and quantitative aspects of the intracellular metabolism of the yeast *Saccharomyces cerevisiae* and will present, describe, and apply tools for the determination of the cells composition, its metabolic rates, and its intracellular concentrations.

**Experimental Tools**

**The common baker’s yeast**

Subject of this study is the yeast *Saccharomyces cerevisiae*, also known under its common name “Baker’s yeast”, a fungi of the order *Saccharomycetales* (budding yeasts) and probably the best characterized eukaryotic organism.

The common baker’s yeast has accompanied humanity since millennia. Its principle metabolic products, carbon dioxide and ethanol, have been employed by humans since the dawn of civilisation for the raising of dough and spirits. Thus, yeast has established itself as being beneficial to humanity and is generally regarded as safe. In recent years, a number of new uses have added to the attractiveness of *S. cerevisiae*, the expression of heterologous proteins, its use as food supplement for feedstock, e.g. as additional lysine source (Niederberger, 1989), and the production of bio-fuel are some examples.

The large scope of different possible applications are one reason for the vast amount of research done on this species. Being a eukaryotic organism, and thus structurally closer to the human cell, yeast often functions as a model organism, in which to investigate phenomena, such as the intracellular targeting of proteins. This is further helped by the relative ease of its cultivation. A high growth rate, low media requirements, a robustness against changes in its environment, high mechanical strength, the ability to grow aerobically as well as anaerobically, and its GRAS (generally recognized as safe) status are among the reasons, which made yeast a favourite in many research laboratories. As a result of its popularity, it was among the first organisms, for which the complete genome was decoded (Goffeau et al., 1996).

The haploid autotroph strain CEN.PK133-7D (*MATα, MAL2-8′ SUC2*), also deposited as CBS8340 at the Dutch “Centraal Bureau voor Schimmelculturen”, was used during the course of this study. Its isogenic diploid CEN.PK122 was chosen from a shortlist of four different strains at the beginning of the project for its accessibility to genetic manipulation and decent performance under fermentation conditions (van Dijken et al., 2000).

**Cultivation Tools**

Development of quantitative methods in microorganisms requires its growth under defined conditions to provide a reproducible reference state against which mathematical models
can be tested. The latter requires the minimisation of possible growth variation. The use of a defined mineral medium excludes naturally occurring fluctuations of complex supplement and allows thus a better characterisation of the system. Another requirement is the exclusion of microorganisms other than the studied species, thus sterile conditions, commonly achieved through heat sterilization of the vessel and either heat or filter sterilization of the growth media.

![Graph showing growth curves for different strains](image)

**Figure 3:** shake flask cultivation of different *S. cerevisiae* strains on mineral medium with glucose as carbon source at T=30°C and 200 rpm; biomass concentration was measured via optical density (OD) at 660 nm. The slope of the normalized biomass concentration plotted on a natural logarithmic scale gives the maximal growth rate.

Four methods are widely used for sterile cultivation of microorganisms. The simplest and most common one is probably the cultivation on plates, where a solid medium, often agar, provides the physical support as well as the nutrients for microbial growth. Cells are smeared onto the surface of the plates in such a manner, that ideally each cell forms a separate colony of daughter cells. Ideally suited for characterizing the organism, especially in respect to growth requirements, this method is mostly used for biological and genetical studies of the microorganism. It is, however, limited by its biomass amount and the ability of monitoring the local growth conditions, which might vary significantly for different cells depending to their position within a colony.

The use of a shake flask cultivation, whereby the microorganism is cultivated in a liquid growth medium and kept suspended through agitation of the vessel, eliminates the local gradients in the culture. The method is applied for physiological studies, the provision of small amounts of biomass, and simple quantitative studies, such as the determination of maximal growth rates (figure 3). While local concentration gradients are minimal, they cannot always be completely excluded due to mass transfer limitations, for example for oxygen. Furthermore, the possibilities to control and measure the growth conditions are limited. Finally, changing nutrient concentrations and pH levels seriously hamper any efforts to obtain a reproducible and well-defined state.
Therefore, the more appropriate tools for the quantitative analysis of microbial growth are stirred tank cultivations in either batch or continuous mode. Microbial growth occurs here in dedicated reactors with control of the physical parameters such as the gas- and liquid flow rates, stirrer speed, pH, temperature, and dissolved oxygen.

**Batch cultivation of yeast**

Growth of yeast in batch fermentation on fermentable substrate proceeds via several distinct phases. After inoculation and a possible short lag phase of about one hour, growth occurs in an environment with practically unlimited resources (figure 4). Multiplication of biomass proceeds at the microorganisms inherent maximal growth rate $\mu_{\text{max}}$ of

![Graph showing growth phases](image)

**Figure 4**: batch cultivation of *S. cerevisiae* BAY17 on mineral medium and glucose as carbon source in a stirred tank reactor. Shown are concentrations of the biomass in the fermentor and of oxygen and carbon dioxide in the off-gas

approximately 0.4 - 0.5 hr$^{-1}$. During this phase of exponential growth, *S. cerevisiae* displays the so called Crabtree effect, where, despite sufficient oxygen, the metabolism is fermentative, producing mainly ethanol from the consumed sugar. Catabolite repression prevents the simultaneous consumption of glucose and ethanol, thus the ethanol accumulates until the glucose is completely depleted, which happens after approximately 18 hours. Towards the end of this phase, an extracellular accumulation of several metabolites can be observed. Typical byproducts are acetate, pyruvate, and glycerol (figure 5). The end of the exponential growth phase is characterized by an almost instant halt of oxygen consumption and carbon dioxide production.

The phase of glucose growth is succeeded by the diauxic shift, where the cell’s metabolism changes from glucose to ethanol consumption. Exchange rates of oxygen and carbondioxide remain on a low level during these period of approximately two hours.
Biomass production practically ceases during the diauxic shift, while the cell’s enzymatic makeup is transformed (figure 4 & 5).

![Graph showing metabolite concentrations over time](image)

**Figure 5:** extracellular metabolite concentrations during batch cultivation of *S. cerevisiae* BAY17 on mineral medium and glucose as carbon source in a stirred tank reactor

Subsequently, the batch culture enters a second phase of intensive biomass production, although the maximum growth rate is much lower than during the initial exponential phase. The produced ethanol is consumed, causing a large increase in the oxygen uptake rate. The period of ethanol consumption lasts for approximately 15 hours, during the end of which the produced acetate is co-consumed, causing an increase of the pH in the fermentation broth unless titrated with acid. The end of the ethanol growth phase is marked by sharp increase in the carbon dioxide evolution rate, which about doubles and subsequently collapsed to near zero values in a time frame of 20 minutes (figure 4).

The length of the different periods largely depends on the condition of the inoculum. An inoculation with a preculture in the exponential growth phase close to its end, eliminates the initial lag phase almost completely. Earlier inoculations will lead to a slower biomass build-up, thus a longer exponential phase; inoculations with older precultures will lead to a longer lag phase, but a start from a higher level of biomass. Furthermore, patterns of metabolite formations can vary sharply between different strains.

Therefore, the reproduction of batch experimental conditions require a very strict time table. Only precise timing, quality, and quantity of inoculation will insure identical results between different experiments. Repetition of an analysis at a certain state of the fermentation requires thus repeating the complete experiment. Additionally, the characterization of the biological state during the analysis demands the complete analysis of the system at the very moment of the analysis, due to the constantly changing nature of the fermentation.
A variation of the batch fermentation is the fed-batch cultivation, whereby after the initial exponential growth period with $\mu_{\text{max}}$, the limiting substrate is added to the fermentor in a controlled fashion to enable optimal formation of biomass and/or product at a rate sustainable by the fermentation set-up. This allows the increase of the biomass concentration beyond the possibilities of a normal batch fermentation, where substrate inhibition or transport limitations of, e.g., heat or oxygen, limits the maximal possible biomass density. Both, the batch and, especially, the fed-batch fermentation are the preferred modes of operation in large scale industrial applications, where a high volumetric productivity is the objective. Despite this, their constantly changing state parameters make these experimental routes not very attractive for quantitative analysis of intracellular phenomena.

![Graph](image)

**Figure 6:** typical stages of an aerobic chemostat cultivation of *S. cerevisiae* CEN.PK113-7D on mineral medium and glucose as carbon source: initial batch phase (A), reaching of steady state (B) and steady state (C)

**Chemostat cultivation of yeast**

The preferred choice of microbial cultivation in the research of quantitative physiology of intracellular metabolism is, therefore, the chemostat fermentation. Here, the constant inflow of growth medium is balanced by an equal outflow of broth, thus keeping a unchanging amount of liquid in the fermentor and resulting in a constant ratio of effluent rate and broth volume, the dilution rate $D$. This leads to the establishment of dynamic steady state, in which all the concentrations in the fermentor are usually assumed to be constant and all fluxes (e.g. $O_2$ uptake and $CO_2$ production) are in steady state, once the culture has passed the initial batch phase and an additional period of about five dilution time (figure 6). The operational mode can be described by the tools developed in the chemical technology for a continuous stirred tank reactor (CSTR).
The biomass withdrawn with the effluent is compensated for by growth in the fermentor, thus, under steady state conditions, the dilution rate $D$ equals the specific growth rate $\mu$ and can be easily manipulated through control of the liquid flow rate. Concentrations in the feed are chosen to support a defined amount of biomass, with one, rarely two substrates being the limiting factor, while all others are provided in excess. This leads to the almost complete consumption of the limiting compound, leaving only residual concentrations in the effluent.

![Graph showing biomass yield and biomass over time](image)

**Figure 7**: changes in biomass concentration and yield for *S. cerevisiae* CEN.PK113-7D grown aerobically on mineral medium during a long term chemostat cultivation, $\mu = 0.1 \text{ hr}^{-1}$, pH = 5.0, T = 30°C with glucose as limiting carbon source

The theory of chemostat cultivation assumes a non-variable behaviour of the cultivated microorganisms, thus e.g. the biomass yield on the limiting substrate is presumed to be a constant factor. However, the low concentration of the limiting substrate poses a strong selective pressure on the cells. As a result, cells with superior traits, e.g. an improved uptake mechanism, will out-compete others. This leads to a gradual shift in the properties of the culture, i.e. a decrease in the residual substrate concentration in the broth (Rutgers et al., 1987) and a simultaneous increase in the biomass yield (figure 7). The change of culturing conditions is also reflected in the intracellular conditions of the cell, for example in a change of intracellular metabolite concentration (figure 8) and changes in the RNA expression patterns (Ferea et al., 1999). This phenomena has been known for a long time, and is exploited for the creation of variant strains with certain properties, for which was selected by applying the appropriate selective pressure. Therefore, a flux steady state does not implicitly guarantee a physiological steady state. A well-defined state of a biological system has thus always to include a time dimension.
Figure 8: changes in measured intracellular concentration (µmol/g DW) in of *S. cerevisiae* CEN.PK113-7D during long term chemostat cultivation at a growth rate of \( \mu = 0.05 \, \text{hr}^{-1} \), pH = 5.0, T = 30°C on a 95:5 mixture of glucose and ethanol. Measurements were performed according to Lange et al. (Lange et al., 2001)

Besides these long term effects, several other observations might disturb a steady state situation in a chemostat. The existence of one steady state is questioned by the mathematic postulate of a multitude of glycolytic steady states in bacteria (Hatzimanikatis and Bailey, 1997a) and the occurrence of metabolic oscillations in e.g. the glycolysis (Boiteux et al., 1975). Clearly detectable are cell cycle oscillations, during which the majority of cells synchronously proceed through the different stages of the cell cycle with dramatic effects on the fermentation as shown in figure 9. It is therefore obvious, that the concept of a steady state in a biological system is an idealistic simplification of the complex behaviour of microorganisms under these conditions. Despite these limitations in a chemostat cultivation, it remains the most suited tool for quantitative analysis of the cell metabolism under well-defined conditions.

Outline of the thesis

This introduction briefly outlined the tools currently available for the investigation into the cell’s metabolism. Some of the possible pitfalls were demonstrated and some open questions pointed out. Based on this knowledge, the following chapters of this thesis are dealing with new tools for the analytical and mathematical quantification of the metabolism of *Saccharomyces cerevisiae* CEN.PK 113-7D and mutants derived thereof. The presented tools and results are a step forward into the direction of a quantitative description of the cell’s enzymatic activities. Thus, they lay the foundation for the future dynamic modelling of the metabolism of yeast.

The second chapter of this thesis analyses the biomass composition of the yeast *S. cerevisiae* under different growth conditions. Various analytical methods for the
determination of the biomass constituents are applied, resulting in a large amount of partly conflicting data. Statistical tools are developed to test and reconcile these different measurements. This allows the quantitative comparison of the different methods, the detection of gross errors in the applied methods, and the identification of omitted compounds. Using this mathematical tool, reconciled biomass compositions are given for a variety of different growth conditions, laying the basis for an improved metabolic flux analysis under these terms.

![Graph](image)

**Figure 9:** autonomous cell cycle oscillation in of *S. cerevisiae* CEN.PK113-7D grown on mineral medium and glucose as carbon source

In the following chapter the predictive capacities of non compartmented models are tested on deletion mutants. A simple metabolic network is employed to initially describe the growth of yeast on glucose, ethanol, and acetate. It is applied to derive the energetic parameters for aerobic growth under carbon-limited chemostat growth from a set of experiments performed over a range of different growth rates and different substrates. The obtained parameters are subsequently used to model the growth of deletion mutants of glyoxylate and gluconeogenic cycle enzymes, which are restricted to an increasing degree in their capacity to metabolise ethanol. The results of the mathematical predictions are finally compared with the measurements from chemostat experiments with these mutants, shedding light on the predictive capacities of the metabolic flux analysis.

To overcome the limitations of a non-compartmented model, the chapter four embarks on the task of giving an overview of the current knowledge of the compartmentation of the metabolism in yeast. Special emphasis is put on the transport proteins, which connect the pathways of different compartments. Based on this overview, a reduced metabolic network is proposed, which includes a cytosolic and a mitochondrial compartment linked via the appropriate transport systems identified for yeast.
The energetic aspects of compartmentation are discussed in chapter five. Based on chemostat experiments of *S. cerevisiae*, the metabolic fluxes and the energetics of the cell are analysed using the compartmented model obtained in the previous chapter. This allows a first quantification of metabolite fluxes through different transport proteins, clearly demonstrating e.g. the advantages of antiport systems employed. Furthermore, the energetic consequences of the transport costs are demonstrated on the differences of consumed energy to regenerate mitochondrial and cytosolic ATP.

The sixth chapter uses the metabolic flux analysis based on cofactor balances and the one based on $^{13}$C-labelling. Comparison of the results of both approaches is performed for steady state chemostat cultivations of a wildtype strain and its isogenic triose phosphate isomerase deletion (*tpiA*) mutant. This approach does not only allow the quantification of fluxes in a given metabolic network, but also the selection of the most likely network from a set of possible metabolic options, thus also providing a tool for qualitative network analysis.

In the final, seventh chapter, the setup of a rapid sampling procedure for the determination of intracellular metabolite concentrations is described. The newly developed method allows the precise determination of intracellular concentrations in yeast, using only very small sampling sizes. The technique gives thus the stepping stone for the next phase of mathematical models of the yeast’s metabolism, the dynamic modelling.

**References**


Small, J.R., Kacser, H. 1993a. Responses of metabolic systems to large changes in enzyme activities and effectors

Small, J.R., Kacser, H. 1993b. Responses of metabolic systems to large changes in enzyme activities and effectors


Chapter 2

Statistical Reconciliation of the Biomass Composition

published as

Statistical Reconciliation of the Elemental and Polymeric Biomass Composition of Saccharomyces cerevisiae
H.C. Lange, J.J. Heijnen

in

Biotechnology and Bioengineering
Vol. 75(3), 334-344, 2001
Statistical Reconciliation of the Biomass Composition

Abstract

A systematic mathematical procedure capable of detecting the presence of a gross error in the measurements and of reconciling connected data sets by using the maximum likelihood principle is applied to the biomass composition data of yeast. The biomass composition of *Saccharomyces cerevisiae* grown in a chemostat under glucose limitation was analysed for its elemental and for its molecular composition. Both descriptions initially resulted in conflicting results concerning the elemental composition, molecular weight, and degrees of reduction. The application of the statistical reconciliation method based on elemental balances and equality relations is used to obtain a consistent biomass composition. Simultaneously the error margins of the data sets are significantly reduced in the reconciliation process. Based on statistical analysis it was found that inclusion of about 4% water in the list of biomass constituents is essential to adequately describe the dry biomass and match both set of measurements. The reconciled carbon content of the biomass varied 4% from the ones obtained from the molecular analysis. The proposed method increases the accuracy of biomass composition data in terms of its elements and its molecules by providing a best estimate based on all available data, and thus provides an improved and consistent basis for metabolic flux analysis as well as black box modelling approaches.
Introduction

Cells consist of a large variety of different biopolymers and macromolecules. Knowledge of their composition and quantity is essential for a metabolic and energetic analysis of the biomass growth. However, analysis is labourious and prone to errors due to the complexity of the biomass. Commonly either of two analyses is performed: the elemental composition is required for the ‘black box’ description of the biomass (Duboc et al., 1995; Heijnen and Roels, 1981b; Verduyn et al., 1991), while analysis of the macromolecules must be considered for flux calculations in metabolic models (Holms, 1986; Parrau et al., 1991). However, Valino and Stephanopoulos (Vallino and Stephanopoulos, 1993) based their flux analysis on the measured elemental composition and the molecular biomass composition was fitted accordingly. Precise knowledge about the biomass composition is required for parameters derived from flux analysis, such as the P/O ratio and the growth-dependent maintenance. Exact data is further needed in dynamic models to, e.g., quantify the depletion of intracellular metabolite pools during pulses due to anabolic fluxes.

Considerable effort has been spent in the exact determination of the elemental composition and the required analytical procedure (Battley, 1995; Duboc et al., 1995; Gurakan et al., 1990; Vriezen, 1998). While C, H, and N determination is commonly done with good precision, the oxygen content is rarely measured, but derived through balancing with the measured elements and ash content (Nielsen and Villadsen, 1994). Battley (Battley, 1995) points out that the quantification of elements through ash measurements can lead to significant errors. Mainly the unknown quantity of oxygen bound in the ash and the unclear fate of sulfur during the ashing process cause a large uncertainty in the ash determination and subsequently in the amount of oxygen in the biomass (Battley, 1995; Duboc et al., 1995). Dried biomass readily and rapidly absorbs humidity changing weight and composition of the sample, thus further distorting the hydrogen and oxygen measurements (Gurakan et al., 1990). Nonetheless the material and energetic balancing of biological processes requires the exact quantity of the latter.

The quantification of the cellular components like lipids or carbohydrates has received much attention (Herbert et al., 1971; Oura, 1972; Schulze, 1995; Stewart, 1975; Vriezen, 1998). For simplicity the biomass is commonly described as consisting of five groups of macromolecules: proteins, carbohydrates, lipids, RNA, and DNA. Together with water and metals these components give the molecular composition of the biomass. The required analytical methods for these macromolecules are labourious and prone to interference. This is due to the wide range of different molecules collectively measured as, e.g., carbohydrates and the varying degree of polymerisation in different proteins. Again, the hygroscopic nature of dried biomass complicates the measurements. Precise quantification of the water content is problematic. Hence, complete analysis of a microorganism grown under well defined growth conditions is rarely reported in terms of both its elemental and its molecular composition.

It is obvious that the biomass composition in terms of its elements as well as its molecules should be consistent, since a known molecular composition directly defines the elemental
composition. Comparison of both values frequently shows discrepancies (Duboc et al., 1995; Vriezen, 1998). Duboc et al. (Duboc et al., 1995) state that this mismatch is mainly due to large errors in the quantification of the polymer content. Schulze (Schulze, 1995) reports the measured C:N ratio together with the molecular biomass composition; the C:N ratio as measured by elemental analysis lies on average 20% above the value derived from the molecular components. However, a matching of the information from both elemental and molecular analysis has so far not been reported.

Nevertheless, the precise knowledge of these values is essential for energetic and metabolic calculations. Vriezen et al. (Vriezen, 1998) note that fluxes in a metabolic network of hybridoma cultures may vary up to 50% depending on the measured polymer composition. Nissen et al. (Nissen et al., 1997) show that already a small error in the protein measurements leads to large errors in the calculated metabolic fluxes, e.g., errors in protein content are magnified by a factor of two for mitochondrial NADH generation.

In this article a consistent description of dry biomass composition of Saccharomyces cerevisiae with respect to elements and polymers is obtained from the original set of measurements using available linear constraints and statistical analysis of the accuracy of the available data. The linear constraints are based on the elemental balances and equality relations obtained through the use of different analytical techniques. Furthermore, the chi-square value of each data set is calculated indicating the quality of agreement between measured and calculated data and disclosing faulty measurements. The reconciled measurements give a consistent description of the biomass with greatly reduced error margins for its elemental and molecular content. Different analytical methods for the carbon, nitrogen and protein determination are applied to improve the quantitative description of the biomass. Different drying methods are compared to obtain the amount of the remaining water in the dried biomass.

Materials and Methods

Strain cultivation
S. cerevisiae CEN.PK113-7D was obtained from the Centraalbureau voor Schimmel-cultures (CBS), The Netherlands; frozen stock solution was kept at -80°C. Chemostat cultivation was performed in 2 litre Applicon fermentors at dilution rates of 0.02 -0.2 hr⁻¹, 30°C and pH 5.0 under carbon-limited conditions on defined mineral media according to Verduyn et al. (Verduyn et al., 1992). Dissolved oxygen tension remained always above 30%. The fermentors were inoculated with approximately 50 ml preculture, grown overnight on the same mineral medium and 10 g/l glucose as carbon-source. Glucose served as sole carbon and energy source, its feed concentration varied for different cultures between 5 and 15 g/l glucose. The ammonia concentration of the feed was 1.29 g/l for glucose concentrations less than 10 g/l and twice this value for higher glucose concentrations. Carbon balances were checked routinely and closed with more than 95% recovery for all cultures.
Biomass analysis

Biomass was harvested after reaching steady state. A culture was considered to be in steady state if the biomass concentration was constant for at least five residence times and the biomass was devoid of cell cycle associated oscillations (Duboc et al., 1996). The cells were spun down and washed with demineralized water. The biomass pellet was stored at -20°C until being freeze-dried for 48 hours at 10 Pa. The biomass was then further dried at 70°C for 48 hours and stored at room temperature in a desiccator above silica gel to obtain a reproducible reference state until further analysis.

Amino acids in the proteins were determined from 10 μl cell suspension with 2 g/l biomass hydrolysed at 110°C for 22 hours in 200 μl solution of 6N HCl and 1% phenol. During hydrolysis the samples were continuously flushed with nitrogen. The hydrolysed samples were vacuum dried and dissolved in distilled water. Amino acid concentration was measured by HPLC using the AccQ-tag system, as described by Vriezen et al. (Vriezen et al., 1997). The method's recovery was determined as 0.82 ± 5% based on measurements of different pure protein and corrections were made accordingly.

The carbohydrates were measured according to Herbert (Herbert et al., 1971): 5 ml concentrated sulfuric acid were added to a mixture of 1 ml sample solution (0.2 mg dry biomass/ml) and 1 ml aqueous phenol solution (50 g/l) (Schulze, 1995). Absorbance at 488 nm was measured after 30 minutes incubation at room temperature using glucose solutions as standard. The results were corrected for the presence of nucleic pentoses using a relative absorbance of 0.445 and 0.264 for RNA and DNA respectively. Based on experiments with different polysaccharides a recovery factor of 0.95 ±7% was applied to the biomass data.

Total protein was determined using the Biuret method. 2 ml resuspended dried biomass (8 g/l) together with 1 ml 1M NaOH was incubated at 100°C for 10 minutes and subsequently cooled on ice. 0.9 ml solution was mixed with 0.3 ml of 0.1M copper sulphate solution, incubated for 5 min at room temperature, and finally centrifuged for 5 minutes. The absorbance of the clear supernatant was measured at 510 nm using dried BSA as standard (Verduyn et al., 1990).

The cellular lipid components of 100 mg biomass dissolved in 100 μl H₂O were extracted twice with 3.75 ml of a 2:1 mixture of methanol and chloroform. Extraction was facilitated by placing the samples in an ultrasonic bath. The extract was washed with 1.25 ml 0.9% NaCl solution; the two phases were separated by centrifugation and decanting. The chloroform was evaporated overnight under a nitrogen atmosphere. The remaining lipid was quantified gravimetrically (Henriksen et al., 1996). Measurements were done in quadruplicate to reduce the noise.

RNA and DNA analyses were performed according the Schmidt-Thannhauser-Schneider method as described by Herbert et al. (Herbert et al., 1971). Absorption was measured at 260 and 280 nm, the latter to give an indication for potential impurities caused by protein. The amount of nucleic acids was calculated as described by Benthin (Benthin et al., 1991). Values were corrected for recovery, which was quantified as 0.9 ±5% with bakers yeast.
RNA (Sigma, R-6750, less than 3% impurities) and 0.8 ±5% with herring sperm DNA (223646, Boehringer Mannheim, Germany). The purity of DNA was unknown and taken as 100%.

The elemental contents of C, H, N, and S were determined as described by Vriezen et al. (Vriezen, 1998). For analysis of P, K, Na, Mg, Ca, Fe, Mn, Cu, and Zn, 0.5 g of biomass was treated in concentrated sulfuric acid overnight. Subsequently the elemental concentrations were measured using inductively coupled plasma (ICP) with argon as plasma gas.

The carbon content was further measured by TOC (Dohrmann DC-190, Rosemount Analytical, Santa Clara, Ca) based on the difference of at least five measurements of fresh fermentation broth and its filtrate. The carbon fraction \( f_{C:TOC} \) was then calculated based on the simultaneous measurement of the dry biomass concentrations in the fermentor. The nitrogen content was derived through balancing the ammonia feed and its outflow in the supernatant of the broth. Ammonia measurements were performed on a Scalar Autoanalyser 5100 measuring absorption at 650 nm of a 1:1:2 mixture of sample, hypochlorite, and salicylate solutions. The biomass concentration was determined through filtration of the broth with a nitrocellulose filter (Satorius, Braunschweig, Germany, pore size 0.45 μm), drying at 70°C for 48 hours and subsequent storage above silica gel.

**Theory of balancing**

**Biomass composition**

Chemical analysis of biomass can be done on two levels resulting in either an elemental or a molecular description of its composition. At the elemental level the biomass is analysed for its C, H, N, O, P, S and metal content, the metals are lumped together in a pseudo element M. At the second level biomass is described in terms of proteins, carbohydrates, lipids, RNA, DNA, phosphate, sulphate, and water. Obviously, both sets of values should be compatible with each other if the elemental composition of the macromolecules is known.

The following premises are made for the macromolecule composition of yeast: Their composition is independent of the growth rate. The carbohydrates can be described as polyhexoses with infinite linear chains. For the nucleic acids a guanine and cytidine content of each 20% is assumed. Lipids are taken as 60% fat and 40% phospholipids, the acyl groups

<table>
<thead>
<tr>
<th>biomass constituents</th>
<th>( \text{CH}<em>{1.581} \text{N}</em>{0.275} \text{O}<em>{0.318} \text{S}</em>{0.003} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>protein(^a)</td>
<td></td>
</tr>
<tr>
<td>carbohydrate(^b)</td>
<td>( \text{CH}<em>{1.667} \text{O}</em>{0.833} )</td>
</tr>
<tr>
<td>lipids(^b)</td>
<td>( \text{CH}<em>{1.873} \text{N}</em>{0.0100} \text{O}<em>{0.140} \text{P}</em>{0.010} )</td>
</tr>
<tr>
<td>RNA(^b)</td>
<td>( \text{CH}<em>{1.232} \text{N}</em>{0.389} \text{O}<em>{0.737} \text{P}</em>{0.105} )</td>
</tr>
<tr>
<td>DNA(^b)</td>
<td>( \text{CH}<em>{1.255} \text{N}</em>{0.378} \text{O}<em>{0.612} \text{P}</em>{0.102} )</td>
</tr>
<tr>
<td>phosphate</td>
<td>( \text{H}_3\text{PO}_4 )</td>
</tr>
<tr>
<td>sulphate</td>
<td>( \text{H}_2\text{SO}_4 )</td>
</tr>
<tr>
<td>water</td>
<td>( \text{H}_2\text{O} )</td>
</tr>
<tr>
<td>metals</td>
<td>M</td>
</tr>
</tbody>
</table>

\(^a\) as derived from amino acid analysis  
\(^b\) according to Kockova-Knatochvilova, 1990

**Table 1:** composition of biomass constituents used for reconciliation
are made up of palmitoleic, oleic, palmitic, and stearic acids in a ratio of 44:17:14:10, as reported for \textit{S. cerevisiae} grown on synthetic media by Kocková (Kocková-Kratochvílová, 1990). One acyl group is substituted in the phospholipids with an average of 23, 62, and 16\% of ethanolamine, choline, and serine respectively bound through a phosphate group. The protein composition was derived from the experimental analysis of the amino acids. The resulting elemental composition of each macromolecule is given in Table I.

\section*{Relationships between measurements}

The biomass is analysed for all its components \( i \), elements and molecules; the results are expressed as mass fraction \( f_i \) of the dry biomass. These results can be related to each other either directly or using mass balances of the elements. The obtained system of linear constraints should be fulfilled by all measurements.

The first and most obvious constraint is obtained through the use of the equivalence of different methods, in our case for the determination of the mass fraction \( f \) of protein, carbon, and nitrogen in the dry biomass. Under the assumption that each method resulted in a valid quantitative description of the measured variable the following equations must hold:

\begin{align}
  f_{\text{protein-biuret}} &= f_{\text{protein-amino acids}} = f_{\text{protein}} \quad (a) \\
  f_{C-EA} &= f_{C-TOC} = f_C \quad (b) \\
  f_{N-EA} &= f_{N-AA} = f_N \quad (c)
\end{align}

The elemental balance results in constraints of the second kind. The measured mass fraction \( f \) of the element \( i \) in the biomass must match the sum of its content \( x_i \) in each molecule, thus linking elemental and molecular description of the biomass:

\begin{equation}
  f_i = x_{i,\text{protein}} \cdot f_{\text{protein}} + x_{i,\text{carb}} \cdot f_{\text{carb}} + x_{i,\text{RNA}} \cdot f_{\text{RNA}} + \ldots (2)
\end{equation}

The elemental content \( x_i \) is specified by the composition of each molecule. Therefore, the system comprises seven elemental restrictions.

The last relation is based on the total mass balance. If all elements and all molecules are included in the model, the sum of the elements as well as macromolecules must comprise the whole biomass. Therefore, the inorganic cations K, Mg, Na, Fe, etc. have to be included to complete the list of elements and are pooled together as one pseudo element ‘M’. Expressed in mass fractions one can write:

\begin{align}
  \sum f_{\text{elements}} &= 1 \quad (a) \\
  \sum f_{\text{molecules}} &= 1 \quad (b)
\end{align}

Since the metals are not accounted for in the organic molecules, a further molecule ‘metals’ is introduced to complete the list of molecules. Equation (3b) can be obtained from linear combination of equation (3a) and (2), thus only one further independent
relation is gained. The whole set of linear constraints can be represented in matrix form by

\[ C \cdot f = b \]  \hspace{1cm} (4)

where \( C \) is the matrix of constraints containing one row containing the weight restraint (3a), six rows of equality relations (1a-c), and seven rows derived from the elemental balances of C, H, N, O, P, S, and M (2). The vector \( f \) with its elements \( f_i \) represents the biomass constituents. The exact structure is given in the appendix.

**System analysis and statistical reconciliation**

It is evident that the mass fractions of the molecules of which the biomass is composed can completely characterize the biomass. Thus, the degree of freedom \( df \) for this system equals the number of molecules minus one due to the total mass balance. Increasing the number of measurements beyond the degree of freedom creates redundant information, which is used for data reconciliation.

The maximum likelihood method is used to obtain the most probable set of reconciled variables \( f_i \) fulfilling the system of linear constraints (4). A weighting factor is provided by the covariance matrix \( P_m \), which for independent, original measurements is a diagonal matrix with the variances of each measurement as entries. Here, interdependencies occur for the carbohydrate measurement, which is correlated with the RNA and DNA levels:

\[ f_{\text{carb}} = f_{\text{reading}} - 0.445 \cdot f_{\text{DNA}} - 0.264 \cdot f_{\text{RNA}} \]  \hspace{1cm} (5)

as well as for the elemental analysis of C, H, N, and S, which are measured simultaneously on the same subsample with mass \( m_s \):

\[ f_c = \frac{m_c}{m_s}, \quad f_H = \frac{m_H}{m_s}, \quad f_N = \frac{m_N}{m_s}, \quad f_S = \frac{m_S}{m_s} \]  \hspace{1cm} (6)

which is subject to a weighing error \( \Delta m_s \) of about 3%. Further errors beyond the variance of the methods are introduced during the determination of the recovery factor for carbohydrate, protein, RNA, and DNA. The covariance matrix was therefore derived from the diagonal covariance matrix of the underlying measurements \( P_{mn} \):

\[ P_m = J \cdot P_{mn} \cdot J^T \]  \hspace{1cm} (7)

with \( J \) being the Jacobian.

The chi\(^2\)-distributed test variable \( h \) was calculated for the reconciled data according to van der Heijden (van der Heijden et al., 1994a) to test for the presence of gross errors. If detected, the vector comparison test was applied to distinguish between wrong system definition and significant measuring errors (van der Heijden et al., 1994b).
The reconciled data fulfills the given system constraints; a reduced error interval can be assigned to each individual measurement; and an inconsistent system definition can be excluded.

**Results**

Analysis of the biomass composition was performed on the seven aerobic glucose-limited cultures 'gluc1' to 'gluc7' of *S. cerevisiae* grown at dilution rates of 0.022 to 0.21 hr\(^{-1}\). Feed concentrations of the glucose varied from 5 to 10 g/l. Analysis for trace elements was performed on biomass harvested from four separated cultures 'gluc8' to 'gluc11' with a glucose feed of 5 to 15 g/l. No significant differences in the biomass yields and compositions were observed for the different cell mass densities under the same dilution rate caused by the different glucose concentrations.

**Definition of dry biomass and its water content**

The basis for all measurements was dried biomass. It is therefore essential to establish a clear, workable and repeatable procedure to obtain 'dry' biomass. A low drying temperature of 70°C and subsequent storage above silica gel was chosen as the standard procedure for obtaining 'dry' biomass for further analysis. Storage in a desiccator conveniently provides a time independent reference state for the biomass. The speed of water absorption was greatly reduced compared with samples taken directly from a drying oven.

<table>
<thead>
<tr>
<th><strong>amino acid composition</strong></th>
<th>Ala</th>
<th>Arg</th>
<th>Asx</th>
<th>Cys</th>
<th>Glx</th>
<th>Gly</th>
<th>His</th>
<th>Ile</th>
<th>Leu</th>
<th>Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9.77</td>
<td>3.86</td>
<td>9.28</td>
<td>0.14</td>
<td>15.48</td>
<td>8.89</td>
<td>1.93</td>
<td>5.89</td>
<td>8.01</td>
<td>6.57</td>
</tr>
<tr>
<td>Met</td>
<td>1.14</td>
<td>0.24</td>
<td>3.76</td>
<td>4.22</td>
<td>5.33</td>
<td>5.57</td>
<td>0.65</td>
<td>1.96</td>
<td>7.33</td>
<td></td>
</tr>
</tbody>
</table>

*Asx = Asp + Asn; Glx = Glu + Gln

**Table II**: amino acid composition of the protein as measured (mol %)

The water remaining in the 'dry biomass' was determined by measuring its weight loss through 24 hour drying at 105°C. The weight loss averaged 3.6%, presumably due to water evaporation. The measurements were hampered by the rapid weight increase after the drying resulting in a standard deviation for the weight difference of 26% between samples. Even drying at 105°C might not remove all water from the biomass (Gurakan et al., 1990) while changes to the extracted lipid were reported at this temperature; evaporation of other volatile compounds can therefore not be excluded.

Analysis of the different methods of pretreatment revealed that drying at 105°C for 24 hours resulted in a 1.5% higher weight loss than drying in a microwave oven or drying at 70°C for 48 hours. This indicated a varying degree of water removal. Nevertheless, this difference was nearly compensated for by the different weight gains during subsequent storage above silica gel. Hence the water content
in the so-called dry biomass is nearly independent of the preceding drying method, if followed by a sufficiently long subsequent storage in a desiccator.

### Measurements of the elemental biomass composition

Dry biomass of the cultures 'gluc1' to 'gluc7' was analysed for its elements C, H, N, O, and S using a CHNOS elemental analyser. The carbon fraction \( f_{C,EA} \) increased from 0.42 to 0.44 with higher dilution rates; \( f_{N,EA} \) increased from 0.062 to 0.08. The hydrogen content \( f_{H,EA} \) and the oxygen content \( f_{O,EA} \) did not exhibit a clear trend and varied around 0.065 and 0.38 respectively (table III). The average sulphur content \( f_{S,EA} \) was 0.0015. The \( \text{SO}_2^- \) signal could frequently not be clearly separated from the water peak causing a high standard deviation; in cases of non-detection of the peak a \( f_{S,EA} \) of 0.0015 was assumed for the balancing.

The biomass carbon content as measured by TOC analysis were on average 2% higher than \( f_{C,EA} \) obtained from the CHNOS analysis (table III). The average nitrogen content measured by Autoanalyser \( f_{N,AA} \) of 0.072 compared well with the 0.073 of the elemental analysis but the trend is reversed.

<table>
<thead>
<tr>
<th>D (hr(^{-1}))</th>
<th>( f_{C,EA} )</th>
<th>( f_{C,TOC} )</th>
<th>( f_{H,EA} )</th>
<th>( f_{N,EA} )</th>
<th>( f_{N,AA} )</th>
<th>( f_{O,EA} )</th>
<th>( f_{P,TE} )</th>
<th>( f_{S,EA} )</th>
<th>( f_{M,TE} )</th>
<th>sum$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gluc 1</td>
<td>0.022</td>
<td>41.8</td>
<td>44.1</td>
<td>6.46</td>
<td>6.12</td>
<td>8.35</td>
<td>37.4</td>
<td>0.97</td>
<td>0.16</td>
<td>2.51</td>
</tr>
<tr>
<td>gluc 2</td>
<td>0.052</td>
<td>43.1</td>
<td>45.7</td>
<td>6.59</td>
<td>6.81</td>
<td>8.63</td>
<td>36.4</td>
<td>0.97</td>
<td>0.16</td>
<td>2.51</td>
</tr>
<tr>
<td>gluc 3</td>
<td>0.087</td>
<td>40.4</td>
<td>47.3</td>
<td>5.99</td>
<td>6.77</td>
<td>7.98</td>
<td>36.1</td>
<td>0.97</td>
<td>0.23</td>
<td>2.51</td>
</tr>
<tr>
<td>gluc 4</td>
<td>0.107</td>
<td>45.0</td>
<td>43.9</td>
<td>6.75</td>
<td>7.96</td>
<td>7.32</td>
<td>36.6</td>
<td>0.97</td>
<td>&lt;0.1</td>
<td>2.51</td>
</tr>
<tr>
<td>gluc 5</td>
<td>0.126</td>
<td>43.0</td>
<td>44.2</td>
<td>6.61</td>
<td>7.46</td>
<td>6.49</td>
<td>37.0</td>
<td>0.97</td>
<td>0.13</td>
<td>2.51</td>
</tr>
<tr>
<td>gluc 6</td>
<td>0.158</td>
<td>43.7</td>
<td>45.2</td>
<td>6.63</td>
<td>7.75</td>
<td>6.93</td>
<td>37.0</td>
<td>0.97</td>
<td>0.22</td>
<td>2.51</td>
</tr>
<tr>
<td>gluc 7</td>
<td>0.211</td>
<td>44.2</td>
<td>46.2</td>
<td>6.68</td>
<td>8.55</td>
<td>6.53</td>
<td>35.9</td>
<td>0.97</td>
<td>0.33</td>
<td>2.51</td>
</tr>
</tbody>
</table>

$^a$ determination in triplicate  
$^b$ as \( 0.15 \) used for calculation when below detection limit  
$^c$ determination from at least 5 samples  
$^d$ with unweighted average of \( f_C \) and \( f_{C,TOC} \)  
$^e$ determined from 4 separate steady states at \( D = 0.1 \) hr\(^{-1}\) (table IV)

### Table III: measured elemental mass fractions \( f_i \) in %

Four separate fermentations produced the biomass for the analysis of the remaining elements, metals and phosphorus (table IV). On average the elements K, Mg, and Na made up 2.29, 0.15, and 0.06% respectively of the biomass. Ca, Fe, Mn, Cu, and Zn accounted for a total of 0.02%. Thus, the total metal content was measured as 2.51% of the biomass. In the further calculations this was lumped together into a pseudo element 'M' with an average atomic weight of 37.2 g/mol and a positive charge of 1.1. The phosphorus content \( f_P \) was measured as about 1% of the dry biomass.

The average recovery for the elements of the biomass was 96.9% calculated with an unweighted average of both carbon and both nitrogen measurements. With the measured, raw data from the elemental analysis the unit carbon formula was calculated as \( \text{CH}_{1.769}\text{N}_{0.146}\text{O}_{0.631}\text{P}_{0.009}\text{S}_{0.00162}\text{M}_{0.019} \) for biomass grown at a dilution rate of around 0.1hr\(^{-1}\); the molecular weight was 26.9 g/C-mol and the degree of reduction \( \gamma \) 4.14.
Measurements of the molecular biomass composition

The relative abundance of each amino acid did not vary between the different cultures and the average measured value (table II) was taken to determine the elemental protein composition as CH_{1.58}N_{0.27}O_{0.31}S_{0.003}. Most measured values were within ±10% of the ones reported by Oura (Oura, 1972), Chistyakova (Chistyakova et al., 1982), and Schulze (Schulze, 1995). A distinction between free amino acids and protein was not attempted. Trials with pure proteins exhibited variations in total recovery due to different susceptibility to hydrolyzation since the ratios of amino acids were in agreement with the literature values. On average 0.82 ± 5% of the protein was recovered, thus measurements of the biomass protein content were adjusted accordingly.

<table>
<thead>
<tr>
<th>glucose in feed (g/L)</th>
<th>$f_Ca$ (%)</th>
<th>$f_Cu$ (%)</th>
<th>$f_Fe$ (%)</th>
<th>$f_K$ (%)</th>
<th>$f_Mg$ (%)</th>
<th>$f_Mn$ (%)</th>
<th>$f_Na$ (%)</th>
<th>$f_Zn$ (%)</th>
<th>$f_{metals}$ (%)</th>
<th>$f_P$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gluc 8</td>
<td>5.0</td>
<td>&lt; 0.002</td>
<td>0.0009</td>
<td>0.0028</td>
<td>2.15</td>
<td>0.18</td>
<td>0.0007</td>
<td>0.09</td>
<td>0.01</td>
<td>2.42</td>
</tr>
<tr>
<td>gluc 9</td>
<td>7.5</td>
<td>0.0036</td>
<td>0.0005</td>
<td>0.0034</td>
<td>2.27</td>
<td>0.13 &lt; 0.0001</td>
<td>0.05</td>
<td>0.01</td>
<td>2.47</td>
<td>0.94</td>
</tr>
<tr>
<td>gluc 10</td>
<td>10.0</td>
<td>0.0077</td>
<td>0.0008</td>
<td>0.0036</td>
<td>2.29</td>
<td>0.14 0.0004</td>
<td>0.05</td>
<td>0.01</td>
<td>2.50</td>
<td>0.94</td>
</tr>
<tr>
<td>gluc 11</td>
<td>15.0</td>
<td>0.0038</td>
<td>0.0007</td>
<td>0.0047</td>
<td>2.44</td>
<td>0.15 0.0004</td>
<td>0.04</td>
<td>0.01</td>
<td>2.65</td>
<td>1.02</td>
</tr>
<tr>
<td>average</td>
<td></td>
<td>0.0050</td>
<td>0.0007</td>
<td>0.0036</td>
<td>2.29</td>
<td>0.15 0.0005</td>
<td>0.06</td>
<td>0.01</td>
<td>2.51</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table IV: measured trace elements in dry biomass

The protein content $f_{protein-A}$ calculated based on the summation of the amino acid analysis increased with the dilution rate from 0.38 to 0.46 (table V). The same trend was obtained for the protein content $f_{protein-B}$ analysed by the Biuret method. Results showed a very good agreement between both methods, $f_{protein-B}$ averaged with 0.422 only slightly lower than $f_{protein-A}$ with 0.425, thus confirming BSA as a suitable standard for a valid quantitative protein determination of the measured yeast protein mixture.

<table>
<thead>
<tr>
<th>D (hr⁻¹)</th>
<th>$f_{protein-A}$</th>
<th>$f_{protein-B}$</th>
<th>$f_{carbohydrates}$</th>
<th>$f_{lipids}$</th>
<th>$f_{RNA}$</th>
<th>$f_{DNA}$</th>
<th>$f_{Pi}$</th>
<th>$f_{SOM}$</th>
<th>$f_{water}$</th>
<th>$f_{metals}$</th>
<th>sum$^d$</th>
<th>sum$^A$</th>
<th>sum$^B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gluc 1</td>
<td>0.022</td>
<td>38.4</td>
<td>32.3</td>
<td>45.4</td>
<td>10.2</td>
<td>4.3</td>
<td>0.5</td>
<td>1.5</td>
<td>0.4</td>
<td>3.6</td>
<td>2.5</td>
<td>103.8</td>
<td>106.9</td>
</tr>
<tr>
<td>gluc 2</td>
<td>0.052</td>
<td>42.0</td>
<td>39.2</td>
<td>42.1</td>
<td>9.4</td>
<td>5.2</td>
<td>0.5</td>
<td>1.5</td>
<td>0.4</td>
<td>3.6</td>
<td>2.5</td>
<td>105.9</td>
<td>107.3</td>
</tr>
<tr>
<td>gluc 3</td>
<td>0.087</td>
<td>42.5</td>
<td>42.5</td>
<td>41.2</td>
<td>7.6</td>
<td>6.2</td>
<td>0.5</td>
<td>1.5</td>
<td>0.4</td>
<td>3.6</td>
<td>2.5</td>
<td>106.1</td>
<td>106.0</td>
</tr>
<tr>
<td>gluc 4</td>
<td>0.107</td>
<td>41.4</td>
<td>45.7</td>
<td>38.2</td>
<td>7.2</td>
<td>7.8</td>
<td>0.4</td>
<td>1.5</td>
<td>0.4</td>
<td>3.6</td>
<td>2.5</td>
<td>105.2</td>
<td>103.0</td>
</tr>
<tr>
<td>gluc 5</td>
<td>0.126</td>
<td>44.5</td>
<td>43.0</td>
<td>38.5</td>
<td>10.1</td>
<td>6.8</td>
<td>0.5</td>
<td>1.5</td>
<td>0.4</td>
<td>3.6</td>
<td>2.5</td>
<td>107.7</td>
<td>108.4</td>
</tr>
<tr>
<td>gluc 6</td>
<td>0.158</td>
<td>44.0</td>
<td>45.2</td>
<td>31.6</td>
<td>7.8</td>
<td>7.0</td>
<td>0.5</td>
<td>1.5</td>
<td>0.4</td>
<td>3.6</td>
<td>2.5</td>
<td>99.5</td>
<td>98.9</td>
</tr>
<tr>
<td>gluc 7</td>
<td>0.211</td>
<td>46.3</td>
<td>45.5</td>
<td>31.9</td>
<td>7.7</td>
<td>7.9</td>
<td>0.5</td>
<td>1.5</td>
<td>0.4</td>
<td>3.6</td>
<td>2.5</td>
<td>102.0</td>
<td>102.4</td>
</tr>
</tbody>
</table>

$^*$ protein determined as sum of amino acids
$^a$ protein determined with biuret method
$^b$ not determined for all, average value was assumed
$^d$ with unweighted average of $f_{protein-A}$ and $f_{protein-B}$

Table V: measured mass fractions $f_i$ of the macromolecules in %

The carbohydrate fraction decreased with the growth rate from 0.45 to 0.32, averaging at 0.39 after correction for nucleic acid content and recovery. The method's recovery depended largely on the chain length of the sample. A recovery of about 90% was attained for polymers as glycogen and starch, mono- and disaccharides were completely recovered. Since the degree of polymerisation was not known for the biomass, an average recovery of 0.95±7% was used.

30
The biomass’ average lipids content was 0.086 as measured by chloroform extraction. The biomass had an RNA content between 0.045 and 0.08 increasing with growth rate. The average DNA content was 0.005 without a clear trend. Cultures for which the DNA content was not measured were assigned the average value.

The water content \( f_{H_2O} \) was taken as 0.036 based on the drying experiments. Phosphate and sulphate contents were not measured but taken from literature. Kockova (Kocková-Kratochvílová, 1990) and Duboc (Duboc et al., 1995) report a phosphorous content of 0.015 and 0.02 respectively in S. cerevisiae grown under similar conditions, of which 40% is inorganic phosphate resulting in an approximate phosphate fraction \( f_p \) of 0.015. The sulphate content \( f_{SO_4} \) was estimated accordingly to 0.004.

The elemental composition of the biomass based on the measured macromolecular composition calculated to \( \text{CH}_{1.728}\text{N}_{0.151}\text{O}_{0.577}\text{P}_{0.011}\text{S}_{0.003}\text{M}_{0.017} \) for the cultures ‘gluc4’ and its replicates. These results clearly conflict with the elemental analysis: the molecular weight of one C-mol biomass calculates to 26.1 g/mol, almost 1 g/mol less than determined by elemental analysis. The carbon fraction calculates to 0.49 based on a summary of molecules including water, while only 0.44 was measured with the elemental analysis. The relative \( f_H \) and \( f_O \) were lower despite the addition of water as separate constituent. Without water the unit carbon formula would be \( \text{CH}_{1.631}\text{N}_{0.151}\text{O}_{0.528}\text{P}_{0.011}\text{S}_{0.003}\text{M}_{0.017} \). The C:N ratio differed for both sets: 5.88 versus 5.69 g/g. The phosphor and the sulfur content were slightly lower when determined through elemental analysis. Despite the inconsistencies between both sets of measurements both compare well with values reported in literature (Duboc et al., 1996; Duboc et al., 1995; Larson et al., 1993; Oura, 1972; Verduyn et al., 1991).

<table>
<thead>
<tr>
<th></th>
<th>single sample</th>
<th>between cultures</th>
<th>used in reconcil.</th>
<th>after reconcil.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{C-\text{EA}} )</td>
<td>0.4</td>
<td>3.1</td>
<td>4.3</td>
<td>1.2</td>
</tr>
<tr>
<td>( f_{H-\text{EA}} )</td>
<td>1.1</td>
<td>4.1</td>
<td>5.1</td>
<td>1.3</td>
</tr>
<tr>
<td>( f_{N-\text{EA}} )</td>
<td>0.5</td>
<td>5.3</td>
<td>6.1</td>
<td>2.9</td>
</tr>
<tr>
<td>( f_{O-\text{EA}} )</td>
<td>0.6</td>
<td>2.6</td>
<td>2.6</td>
<td>1.6</td>
</tr>
<tr>
<td>( f_p^a )</td>
<td>-</td>
<td>12</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>( f_{S-\text{EA}} )</td>
<td>15</td>
<td>46</td>
<td>46</td>
<td>21</td>
</tr>
<tr>
<td>( f_{M}^a )</td>
<td>-</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>( f_{C-\text{TOC}} )</td>
<td>2.0</td>
<td>5.8</td>
<td>5.8</td>
<td>1.2</td>
</tr>
<tr>
<td>( f_{N-\text{AA}} )</td>
<td>3.3</td>
<td>7.6</td>
<td>15</td>
<td>2.9</td>
</tr>
<tr>
<td>( f_{\text{protein A}} )</td>
<td>3.2</td>
<td>2.9</td>
<td>5.8</td>
<td>3.4</td>
</tr>
<tr>
<td>( f_{\text{protein B}} )</td>
<td>2.4</td>
<td>5.6</td>
<td>7.5</td>
<td>3.4</td>
</tr>
<tr>
<td>( f_{\text{carbohydr.}} )</td>
<td>3.7</td>
<td>4.3</td>
<td>8.5</td>
<td>4.7</td>
</tr>
<tr>
<td>( f_{\text{lipids}} )</td>
<td>19</td>
<td>20</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>( f_{\text{RNA}} )</td>
<td>12</td>
<td>8.9</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>( f_{\text{DNA}} )</td>
<td>7.1</td>
<td>12</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>( f_{P_i}^b )</td>
<td>-</td>
<td>50</td>
<td>50</td>
<td>41</td>
</tr>
<tr>
<td>( f_{SO_4}^b )</td>
<td>-</td>
<td>50</td>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td>( f_{\text{water}}^c )</td>
<td>13</td>
<td>26</td>
<td>26</td>
<td>19</td>
</tr>
</tbody>
</table>

\( a \) determined only on four separate cultures  
\( b \) estimated  
\( c \) determined separately

**Table VI:** standard deviation of measurements in % of values

The electron balance expressed in the degree of reduction \( \gamma \) can be introduced as a further indicator for inconsistencies within the data. The degree of reduction was calculated with ammonia, phosphate, and sulfate as N, P, and S source respectively (Heijnen, 1981a) as 4.12 when based on the elemental composition and as 4.23 when the composition derived from the polymer analysis was taken.
Chapter 2

Error analysis

The accuracy of the measurements was determined as the standard deviation for multiple analyses of the same sample (table VI, column 1) and as the deviation between samples obtained from different chemostats (table VI, column 2). For the latter, the culturing and analyses was repeated six times at \( D = 0.1 \text{ hr}^{-1} \) (data not shown) to quantify deviations outside the chemical analysis, e.g. culturing conditions, sampling, and sample handling. As expected, deviations between different cultures were larger than on a single sample and formed the basis for the reconciliation. Several measurements were interdependent or were also subject to errors during the quantification of the recovery, the overall error margin as found as entry on the diagonal of the covariance matrix increased for these measurements (table VI, column 3).

The elemental analysis of a single sample varied never more than 1.1%, however, variations between different cultures reached about 5%. Similarly, the higher numbers of samples per culture gave initially a good repeatability of ±2% for the C and N measurements by TOC and Autoanalyser; this despite the subtraction of two values with similar magnitude in the case of the nitrogen content. Comparison between different runs at \( D = 0.1 \text{ hr}^{-1} \) showed almost three-fold higher standard deviations. Of the macromolecules only protein and carbohydrates were determined with a standard deviation of less than 5%, all other values varied with 10% and more. Again, variations between the cultures were higher than on a single sample, with the exception of \( f_{\text{protein}} \) (table VI). The observed differences can only be partially attributed to small variations in the growth rate during the six repetitions of ‘gluc 4’ and thus clearly demonstrate the need to take the sampling procedure into account when determining the accuracy.

Neither the elemental nor the molecular measurement sets summed up to 100% of the biomass: elements added up to about 97% and molecules to 107%. The consistency of these sums indicated a systematic error in some of the measurements.

To test the influence of the error margins on the outcome of the balancing, their size was varied for several biomass constituents. A 10% increase for the standard deviation of RNA and DNA led to a maximum change of 2.5% for the reconciled phosphate levels, RNA and DNA levels changed by around 1%, all other values changed less than 0.5%. Reconciled error margins exhibited a similar trend.

Consistency check and reconciliation

All measurement sets, each describing the biomass of one culture, were rejected with more than 99% certainty by the chi-square test for a biomass lacking water. Serial elimination of individual measurements as proposed by Wang (Wang and Stephanopoulos, 1983) showed a strong negative influence of the nitrogen content \( f_{\text{N,AA}} \) on the chi²-test variable \( h \), but failed to reach statistical acceptance. The vector comparison test as suggested by van der Heijden et al. (van der Heijden et al., 1994b) confirmed that the error could not be related to any of the measurements, but was due to wrong system definition. Levels of \( h \) remained unacceptably high for most cultures after inclusion of water in the list of biomass constituents, but the vector comparison test revealed a good match between the correction and the nitrogen content \( f_{\text{N,AA}} \). Investigations could not find a systematic problem with the
Autoanalyser results, it was therefore concluded that the standard deviation based on the 6 measurements was too low. Statistical acceptance was reached for all cultures after the standard deviation of the nitrogen measurement had been doubled from 7.6 to 15%.

The composition of the biomass constituents was varied to test their influence on the reconciled biomass composition. Each of the elements was individually changed by 5% and the effect on the outcome compared with the original results. Only small changes of 1.5% or less were found for most reconciled values when the protein and carbohydrate composition was varied. However, the protein $f_{\text{protein}}$, the phosphate $f_{\text{P}}$ and the water content $f_{\text{H}_2\text{O}}$ exhibited fluctuations of up to 5.6%. This must be seen as a result of the large uncertainty attributed to these values and is still well within their reconciled error margins. Variations in the composition of the lipids, RNA, or DNA showed similar trends, but a much less pronounced influence on the reconciled data, mostly well below 1%.

Reconciled biomass composition

The average elemental composition of one C-mol biomass after reconciliation was $\text{CH}_{1.48}\text{N}_{0.148}\text{O}_{0.595}\text{P}_{0.006}\text{S}_{0.0019}\text{M}_{0.018}$ for biomass at a dilution rate of $D = 0.1\text{hr}^{-1}$. Accordingly the molecular weight shifted to $M = 26.4\text{g/mol}$.

All elemental contents with the exception of oxygen increased compared to the measured values, the largest relative gain occurred in the nitrogen fraction $f_{\text{N}}$, which increased on average from 0.073 to 0.079. The carbon fraction $f_{\text{C}}$ rose from the measured $f_{\text{C,E4}}$ of 0.432 to 0.455. The hydrogen content $f_{\text{H}}$ increased from 0.065 to 0.067. Oxygen dropped slightly from 0.367 to 0.362. The larger fluctuations in the P and S content are due to the relatively large error of their initial determinations, the reconciled values were 0.01 and 0.0023 for $f_{\text{P}}$ and $f_{\text{S}}$, respectively (table VII). This resulted in a reconciled C:N ratio of 5.76 g/g and a H:O ratio of 0.185 g/g, and thus in a degree of reduction $\gamma$ of 4.19.

<table>
<thead>
<tr>
<th>D (hr$^{-1}$)</th>
<th>$f_{\text{C}}$</th>
<th>$f_{\text{H}}$</th>
<th>$f_{\text{N}}$</th>
<th>$f_{\text{O}}$</th>
<th>$f_{\text{P}}$</th>
<th>$f_{\text{S}}$</th>
<th>$f_{\text{M}}$</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>gluc 1</td>
<td>0.022</td>
<td>45.6</td>
<td>6.8</td>
<td>6.6</td>
<td>37.3</td>
<td>1.00</td>
<td>0.22</td>
<td>2.56</td>
</tr>
<tr>
<td>gluc 2</td>
<td>0.052</td>
<td>46.0</td>
<td>6.8</td>
<td>7.4</td>
<td>36.1</td>
<td>1.00</td>
<td>0.22</td>
<td>2.51</td>
</tr>
<tr>
<td>gluc 3</td>
<td>0.087</td>
<td>45.6</td>
<td>6.7</td>
<td>7.7</td>
<td>36.1</td>
<td>1.02</td>
<td>0.27</td>
<td>2.58</td>
</tr>
<tr>
<td>gluc 4</td>
<td>0.107</td>
<td>45.7</td>
<td>6.7</td>
<td>8.1</td>
<td>35.8</td>
<td>1.03</td>
<td>0.22</td>
<td>2.47</td>
</tr>
<tr>
<td>gluc 5</td>
<td>0.126</td>
<td>45.7</td>
<td>6.8</td>
<td>7.9</td>
<td>35.9</td>
<td>1.03</td>
<td>0.20</td>
<td>2.48</td>
</tr>
<tr>
<td>gluc 6</td>
<td>0.158</td>
<td>45.6</td>
<td>6.8</td>
<td>8.3</td>
<td>35.5</td>
<td>1.04</td>
<td>0.28</td>
<td>2.50</td>
</tr>
<tr>
<td>gluc 7</td>
<td>0.211</td>
<td>45.9</td>
<td>6.8</td>
<td>8.7</td>
<td>34.9</td>
<td>1.04</td>
<td>0.31</td>
<td>2.48</td>
</tr>
</tbody>
</table>

Table VII: balanced elemental mass fractions $f_i$ in %

The expected positive correlation with the dilution rate found in the elemental nitrogen content $f_{\text{N,E4}}$ was confirmed through the balancing procedure. Phosphorous and sulfur levels also increased with the growth rate. The reconciled data revealed a clearly declining trend with increasing growth rates from 0.37 to 0.35 for the oxygen content, which was not evident on the measured data. Values for $f_{\text{C}}$ and $f_{\text{H}}$ remained constant over the whole range of growth rates.
Both, protein and carbohydrate fractions are on average just above 0.39 of the reconciled biomass, with protein increasing from 0.35 to 0.44 while carbohydrates dropping from 0.45 to 0.34. While the carbohydrate content was measured correctly, the protein content was originally overestimated by 6% by the amino acid analysis and the Biuret method. This corresponded to a recovery factor of 0.87 instead of the utilized 0.82 for the sum of amino acids (table VIII).

The reconciled lipid content was on average only 82% of the measured value. This could indicate, that frequently impurities other than lipids were present in the chloroform phase. RNA and DNA amounts were also overestimated. Reconciled mass fractions of RNA increased from 0.04 to 0.07 over the range of dilution rates, $f_{DNA}$ remained constant at 0.0042. This matched a recovery factor of about 1.0 while 0.9 and 0.8 were used for RNA and DNA respectively. The water fraction increased during reconciliation to 0.042, while phosphate and sulphate levels decreased to 0.009 and 0.003 respectively (table VIII).

<table>
<thead>
<tr>
<th>D (hr⁻¹)</th>
<th>$f_{protein}$</th>
<th>$f_{carbohydr}$</th>
<th>$f_{lipids}$</th>
<th>$f_{RNA}$</th>
<th>$f_{DNA}$</th>
<th>$f_{Pi}$</th>
<th>$f_{SMA}$</th>
<th>$f_{water}$</th>
<th>$f_{metals}$</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>gluc 1</td>
<td>0.022</td>
<td>34.5</td>
<td>44.9</td>
<td>8.0</td>
<td>3.8</td>
<td>0.4</td>
<td>1.4</td>
<td>0.3</td>
<td>4.1</td>
<td>2.6</td>
</tr>
<tr>
<td>gluc 2</td>
<td>0.052</td>
<td>38.5</td>
<td>40.9</td>
<td>7.8</td>
<td>4.6</td>
<td>0.4</td>
<td>1.2</td>
<td>0.2</td>
<td>3.9</td>
<td>2.5</td>
</tr>
<tr>
<td>gluc 3</td>
<td>0.087</td>
<td>39.7</td>
<td>40.6</td>
<td>6.2</td>
<td>5.4</td>
<td>0.4</td>
<td>1.1</td>
<td>0.4</td>
<td>3.7</td>
<td>2.6</td>
</tr>
<tr>
<td>gluc 4</td>
<td>0.107</td>
<td>40.7</td>
<td>38.7</td>
<td>6.1</td>
<td>6.6</td>
<td>0.3</td>
<td>0.8</td>
<td>0.2</td>
<td>4.1</td>
<td>2.5</td>
</tr>
<tr>
<td>gluc 5</td>
<td>0.126</td>
<td>40.1</td>
<td>38.8</td>
<td>7.0</td>
<td>5.7</td>
<td>0.4</td>
<td>1.0</td>
<td>0.2</td>
<td>4.4</td>
<td>2.5</td>
</tr>
<tr>
<td>gluc 6</td>
<td>0.158</td>
<td>42.2</td>
<td>35.9</td>
<td>7.0</td>
<td>5.9</td>
<td>0.4</td>
<td>0.9</td>
<td>0.4</td>
<td>4.8</td>
<td>2.5</td>
</tr>
<tr>
<td>gluc 7</td>
<td>0.211</td>
<td>43.8</td>
<td>34.4</td>
<td>6.7</td>
<td>6.6</td>
<td>0.4</td>
<td>0.7</td>
<td>0.4</td>
<td>4.4</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Table VIII: balanced mass fractions $f$, of the macromolecules in %

The error margins were recalculated during reconciliation of the measured data and are reported in the last column of table VI. As expected, the relative error decreases for most measurements. The improvement was most pronounced for the elemental mass fractions with the exception of phosphorus and ‘M’, the protein fraction, and the carbohydrate fraction. Here, the error margins were more than halved. The quality of the quantification also improved albeit less pronounced for phosphorus, ‘M’, lipid, phosphate, and water fractions. The small increases in the relative standard deviations for RNA, DNA, and sulphate could be attributed to the large decrease in their contribution to the biomass, which was not matched by a similar decrease in their absolute error margins.

**Discussion**

A biomass description in terms of its molecular composition is necessarily a compromise between the analytically feasible and the desired precision of the model for which the data is collected. The selected analytical procedures were accepted methods in yeast research and had been employed before. The chosen list of molecules included all essential building blocks of the yeast cell and gave a good starting point for metabolic flux analysis and determination of the cell's energetics. Results from the system consistency check
underlined the assumption that the proposed set of macromolecules was generally valid for analysed yeast biomass.

If raw biomass data was used for macroscopic balancing of fermentations, carbon content and degree of reduction would have been of most importance. The overall carbon balance would show a 1% gap, about 5% additional carbon would be found based on the measured molecular analysis. The latter might easily lead to the omission of a byproduct. While the average of both values was close to the reconciled data for the carbon content, the degree of reduction was too low for both. Parameters calculated from the degree of reduction such as oxygen consumption and CO₂ production would therefore be underestimated unless the reconciled data were taken (table IX).

<table>
<thead>
<tr>
<th>based on:</th>
<th>elemental analysis</th>
<th>molecular analysis</th>
<th>reconciled data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>f_C (%)</td>
<td>γ</td>
<td>f_C (%)</td>
</tr>
<tr>
<td>gluc 1</td>
<td>41.8</td>
<td>4.19</td>
<td>48.0</td>
</tr>
<tr>
<td>gluc 2</td>
<td>43.1</td>
<td>4.22</td>
<td>49.1</td>
</tr>
<tr>
<td>gluc 3</td>
<td>40.4</td>
<td>4.07</td>
<td>48.7</td>
</tr>
<tr>
<td>gluc 4</td>
<td>45.0</td>
<td>4.18</td>
<td>48.2</td>
</tr>
<tr>
<td>gluc 5</td>
<td>43.0</td>
<td>4.16</td>
<td>50.2</td>
</tr>
<tr>
<td>gluc 6</td>
<td>43.7</td>
<td>4.15</td>
<td>46.0</td>
</tr>
<tr>
<td>gluc 7</td>
<td>44.2</td>
<td>4.16</td>
<td>47.1</td>
</tr>
</tbody>
</table>

**Table IX:** carbon content f_C and degree of reduction γ of the measured and reconciled biomass

The measured elemental composition conflicted in several points with the one expected based on the macromolecule content of the biomass. Especially the measured carbon fraction is too low when compared to the expected values based on protein, lipids, etc. Repeated measurements and critical analysis of the sample handling and analysis, did not provide a satisfactory explanation. TOC measurements reinforced the findings of the elemental analysis. Values for elemental analysis as well as molecular composition were consistent with values reported by, e.g., Duboc (Duboc et al., 1996; Duboc et al., 1995), Dekkers (Dekkers et al., 1981), and Schulze (Schulze, 1995) for *S. cerevisiae*. To our knowledge, the discrepancy between these measurements has never been addressed. A possible solution was the introduction of a further molecule with low C-content. Introducing water to the list of compounds was an obvious choice and also mitigated the dilemma of the high oxygen and hydrogen content measured in the elemental composition of the biomass.

The presence of a significant amount of water in the biomass was supported by the statistical analysis, where data for biomass without water was rejected. Experiments showed that dry biomass rapidly absorbs water even in environments considered dry as shown in the example of silica gel. Precise determination of the water was cumbersome. The comparison of biomass weight directly after drying at 105°C with the weight after storage above silica gel was prone to errors and gave only a lower boundary for the water content, since even under ideal conditions more tightly bound water, e.g., on protein
surfaces would not be accounted for. Nevertheless, the method gave a good initial estimate of the water content in the dry biomass of *S. cerevisiae*. The value for the water content agreed with values reported by Larson et al. (Larson et al., 1993). Improved analytical methods for its quantification are desirable, since precise knowledge about the water content is essential for almost any calculation of cell energetics.

Comparison of the TOC measurements with the elemental analysis showed a good agreement between both (table III) despite large variations on individual cultures. The advantages of the TOC measurement performed directly on a liquid sample were the low costs, employment of standard analytical techniques, and the simple sample preparation compared to the elemental analysis. Together with a standard dry biomass determination from the culture a result could be obtained rapidly in most analytical laboratories. A comparison of the standard deviations (table VI) shows that the accuracy of this method must be improved before it can be considered as full substitute for the elemental analysis of carbon. Determination of the nitrogen content via Autoanalyser failed the expectations; its error margins were higher than initially determined.

A common error source for all macromolecular measurements was the recovery factor. While the recovery of carbohydrates was correct, all other ones were found to be too low. This might have been partially caused by lower than assumed purity of the standards, e.g., for RNA and DNA, or incomplete drying in the case of proteins. But it also raises the question of the appropriateness of certain standards, e.g., BSA for the Biuret method. While statistical reconciliation combined with gross error tests could not give a definite answer, it nevertheless provided an insight on the magnitude of the resulting error.

Error margins (table VI) were rather large when compared to values commonly reported for the applied methods (Duboc et al., 1995). But the difference between the standard deviations of measurements on a single sample and between samples of repeated cultures indicated that these values are subject to errors outside the chemical analysis. Sampling and sample handling is a probable source, but small variations in culture conditions are also likely to have an influence. All factors have to be taken into account when assessing the quality of a measurement. The influence of the error margin on the reconciled data is not very strong, but the examples of $f_{N-\text{AA}}$ and $f_{\text{lipid}}$ showed that faulty error estimates might lead to wrong conclusions about the statistical acceptance of the system.

Balancing of the measured values ensured that the reconciled values of elemental and molecular descriptions of the biomass were inherently consistent. The statistical acceptance of the measured data using the proposed molecular compositions indicated a good agreement of model and the real biomass. The general quality of the data was strongly enhanced, as could be seen from the decreased error margins. The resulting data represents the best estimate of elemental and molecular biomass composition based on all available measurements. Its use for black box balancing or metabolic model calculations minimizes errors in, e.g., carbon recovery or calculation of the cell’s energetic requirements.
Acknowledgments

Isabel Campos and Kjell Bangma helped with fermentation and its analysis. Corry Erkelens, Michael Eman, and Gert v.d. Steen carried out the amino acid and the elemental analysis. Hans van Dijken and Nienke Vriezen contributed ideas and constructive criticism. The work is part of the project ‘From gene to product in yeast: a quantitative approach’ and was financed by the European Community through the Framework IV Program on Cell Factories.

List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Autoanalyser</td>
</tr>
<tr>
<td>$D$</td>
<td>dilution rate (hr$^{-1}$)</td>
</tr>
<tr>
<td>$df$</td>
<td>degree of freedom</td>
</tr>
<tr>
<td>EA</td>
<td>elemental analysis</td>
</tr>
<tr>
<td>$f_{i}$-method</td>
<td>mass fraction of element in macromolecule i determined with given method</td>
</tr>
<tr>
<td>$f_{i, balanced}$</td>
<td>mass fraction of i after balancing</td>
</tr>
<tr>
<td>$h$</td>
<td>chi-square test variable</td>
</tr>
<tr>
<td>$J$</td>
<td>Jacobian matrix</td>
</tr>
<tr>
<td>$m_i$</td>
<td>measured mass of component $i$</td>
</tr>
<tr>
<td>$M$</td>
<td>pseudo element ‘metal’ combining all metal ions</td>
</tr>
<tr>
<td>$M_i$</td>
<td>molar mass of compound $i$ (g/mol)</td>
</tr>
<tr>
<td>$P_i$</td>
<td>phosphate</td>
</tr>
<tr>
<td>$P_{E}$</td>
<td>covariance matrix of the residual vector</td>
</tr>
<tr>
<td>$P_m$</td>
<td>covariance matrix of measurements</td>
</tr>
<tr>
<td>$P_{nu}$</td>
<td>covariance matrix of the underlying measurements</td>
</tr>
<tr>
<td>TE</td>
<td>trace element analysis</td>
</tr>
<tr>
<td>$x_{ij}$</td>
<td>mass fraction of element $i$ in compound $j$</td>
</tr>
<tr>
<td>$\delta$</td>
<td>vector of measuring errors</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>degree of reduction</td>
</tr>
</tbody>
</table>

Superscripts

$T$ transpose of a matrix
$^*$ reconciled value
$-$ measured value
Appendix

The linear constraints are represented by the inhomogeneous system

\[ C \cdot f = b \]

with

\[
C = \begin{bmatrix}
C_{11} & C_{12} & C_{13} \\
0 & C_{22} & C_{23} \\
0 & C_{32} & 0
\end{bmatrix}, \quad f = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\]

with the vector of variables \( f \) consisting of

\[
f_1 = \begin{bmatrix}
f_{C-\text{EA}} \\
f_{C-\text{TOC}} \\
f_N \\
f_{N-\text{EA}} \\
f_{N-\text{AA}} \\
f_{\text{protein-\text{amino acids}}} \\
f_{\text{protein-\text{hierac}}}
\end{bmatrix}, \quad f_2 = \begin{bmatrix}
f_C \\
f_{N-\text{EA}} \\
f_N \\
f_{P-\text{TE}} \\
f_{S-\text{EA}} \\
f_{H-\text{TE}}
\end{bmatrix}, \quad f_3 = \begin{bmatrix}
f_{\text{protein}} \\
f_{\text{carbohydrates-phenol}} \\
f_{\text{lipids-extraction}} \\
f_{\text{RNA-SST}} \\
f_{\text{DNA-SST}} \\
f_{\text{PIL-literture}} \\
f_{\text{SO4-literture}} \\
f_{\text{H2O-weight}} \\
f_{\text{water}}
\end{bmatrix}
\]

where \( f_i \) represents fractions of biomass constituents which were measured with more than one methods, all of which were determined. \( f_2 \) and \( f_1 \) represents mass fraction of the elements and molecules respectively which form the biomass, in both cases all but 2 variables were determined. The first row of the submatrices \( C_{1i} \) reflects the equality retrains

\[
C_{11} = I, \quad C_{12} = \begin{bmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}, \quad C_{13} = \begin{bmatrix}
0 & 0 \\
\vdots & \vdots \\
0 & 0 \\
-1 & 0 & \cdots \\
-1 & 0 & \cdots & 0
\end{bmatrix}
\]
The second row of submatrices $C_{2i}$ hold the elemental constrains:

\[
C_{22} = -I, \quad C_{32} = \begin{bmatrix}
X_{C,\text{protein}} & X_{C,\text{carbohydrates}} & \cdots & X_{C,\text{metals}} \\
X_{H,\text{protein}} & X_{H,\text{carbohydrates}} & \cdots \\
\vdots & \vdots & \ddots & \vdots \\
X_{M,\text{protein}} & \cdots & \cdots & X_{M,\text{metals}}
\end{bmatrix}
\]  \tag{12}

and the last row contains the weight restriction

\[
C_{32} = \begin{bmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}
\]  \tag{13}

Separation of the system into a measured and an unmeasured part according to van der Heijden (van der Heijden et al., 1994a) gives

\[
C_m \cdot f_m + C_c \cdot f_c = b
\]  \tag{14}

with the vector $f_c$ containing the unmeasured variables $f_{C_\text{a}}, f_{N_\text{a}}, f_{\text{protein}}$, and $f_{\text{metal}}$. For the error containing vector $\overline{f}_m$ of measured mass fractions holds

\[
\overline{f}_m = f_m - \delta
\]  \tag{15}

with $\delta$ being the vector of measuring errors. The most likely estimate $\hat{\delta}$ of the error vector $\delta$ is obtained from minimization of the objective function

\[
O = \delta^T \cdot P_m^{-1} \cdot \delta
\]  \tag{16}

under the constraint

\[
C_m \cdot \overline{f}_m + C_m \cdot \delta + C_c \cdot f_c = b
\]  \tag{17}

Using Lagrange multipliers one obtains the best estimate for the error vector

\[
\hat{\delta} = P_m \cdot C_m^T \cdot P_{E}^{-1} \cdot (b - C_m \cdot \overline{f}_m - C_c \cdot \hat{f}_c)
\]  \tag{18}

with

\[
\hat{f}_c = (C_c^T \cdot P_{E}^{-1} \cdot C_c) \cdot C_c^T \cdot P_{E}^{-1} \cdot (b - C_m \cdot \overline{f}_m)
\]  \tag{19}
and the covariance matrix of the residual vector defined as $P_E = C_m \cdot P_m \cdot C_m^T$. Thus, the reconciled measurement vector $\hat{f}_m$ is calculated from

$$\hat{f}_m = \overline{f}_m - \hat{\delta}$$

Finally, the reconciled covariance matrix $\hat{P}_m = E[(f_m - \hat{f}_m)(f_m - \hat{f}_m)^T]$ calculates to

$$\hat{P}_m = P_m \cdot \left( I - C_m^T \cdot P_E^{-1} \cdot \left( I + C_c \cdot \left((C_c^T \cdot P_E^{-1} \cdot C_c)^{-1}\right)^T \cdot C_c^T \cdot P_E \right) \right) \cdot C_m \cdot P_m$$

References


Characterisation of null mutants of glyoxylate cycle and gluconeogenic enzymes

published as

Characterisation of null mutants of glyoxylate cycle and gluconeogenic enzymes in *S. cerevisiae* through metabolic network modelling verified by chemostat cultivation.

I. Stückrath, H.C. Lange, P. Kötter, W. van Gulik, K.-D. Entian, J.J. Heijnen

in

*Biotechnology and Bioengineering*

Vol. 77(1), 61-72, 2002
Characterisation of null mutants of glyoxylate cycle and gluconeogenic enzymes

Abstract

Biomass yields for several null mutants in *Saccharomyces cerevisiae* were successfully predicted with a metabolic network model. Energetic parameters of the model were obtained from growth data in C-limited aerobic chemostat cultures of the corresponding wild type strain, which exhibited a $P/O$ ratio of 1.46, a non-growth related maintenance of 56 mmol ATP/C-mol biomass/hr and a growth related requirement of 655 mmol ATP/C-mol biomass. Biomass yield and carbon uptake rates were modelled for different mutants incapacitated in their glyoxylate cycle and their gluconeogenesis. Biomass yields were calculated for different feed ratios of glucose to ethanol, its decrease for higher ethanol fractions were correctly predicted for mutants with deletions of the malate synthase, the isocitrate lyase, or the phosphoenolpyruvate carboxykinase. The growth of the fructose-1,6-bisphosphatase deletion mutant was anticipated less accurate, but the tendency was modelled correctly.
Chapter 3

Introduction

Advances in genetical engineering allow the specific alteration of the metabolic reactions and pathways, but due to the complexity of metabolic networks the outcome of these changes can be predicted intuitively in only few simple cases, mainly in applications concerning a specific production pathway of a secondary metabolite. Directed genetic alterations of the cell's central metabolism require a more systematic approach (Bailey, 1991). Here, metabolic network analysis proved itself to be a valuable tool in predicting the behaviour of the reaction network under various conditions (Holms, 1996; Christensen and Nielsen, 1999).

Metabolic network analysis examines the structure of a proposed metabolic network, thus identifying e.g. parallel pathways, futile cycles, and conserved moieties. It further determines which reactions of a given network can be quantified for a chosen set of measured rates under standard conditions. Metabolic flux analysis is then applied to calculate the rates of these observable reactions. Depending on the structure of the network both tools can be used to obtain the stoichiometrically conceivable yield of a desired product based solely on a set of biological feasible reactions (Vallino and Stephanopoulos, 1993), or its thermodynamically feasible yield by including energetic restrictions through the inclusion of the cofactor balances for NAD(P) and adenosine phosphates (van Gulik and Heijnen, 1995; van Gulik et al., 2001). The availability of sufficient data on a variety of steady state conditions allows the quantification of the cell's energetic parameters; the P/O ratio and the maintenance requirements as demonstrated by Gulik and Heijnen (van Gulik and Heijnen, 1995) and Vanrolleghem et al. (Vanrolleghem et al., 1996).

The yeast *Saccharomyces cerevisiae* is able to metabolize glucose as well as ethanol. Despite the diauxic growth in batch culture *S. cerevisiae* can use glucose and ethanol in mixed cultures simultaneously under carbon-limited conditions in continuous chemostat cultures. Depending on the ratio of glucose and ethanol in the media the activity of the glyoxylate cycle enzymes isocitrate lyase (ICL) and malate synthase (MLS) and the gluconeogenic enzymes phosphoenolpyruvate carboxykinase (PCK) and fructose-1,6-bisphosphatase (FBP) is needed at different values. It has been shown that the onset of these enzymes is tightly regulated depending on the ratio of glucose and ethanol in the substrate (de Jong-Gubbels et al., 1995). By using single null mutants in the mentioned enzymes we could show the ratio of glucose and ethanol in the feed from which the activity of the single enzyme is needed for the complete utilization of both C-sources.

Vanrolleghem et al. (Vanrolleghem et al., 1996) as well as van Gulik and Heijnen (van Gulik and Heijnen, 1995) successfully used metabolic flux analysis to calculate the switching points of these key enzymes. In the presented work, we have now employed metabolic flux analysis for the quantitative prediction of genetically modified *S. cerevisiae* based on knowledge about the unaltered strain. Biomass yields in carbon-limited chemostat fermentations were correctly forecasted over a range of different substrate mixtures based solely on network analysis and energetic parameters obtained from the corresponding wild type strain for different null mutants, incapacitated in their glyoxylate
cycle and their gluconeogenesis. Model predictions were verified with measured data obtained from chemostat cultures.

Materials and Methods

Construction of null mutants

The haploid prototrophic wild type strain *S. cerevisiae* CEN.PK 113-7D was used as reference for the derived knockout mutants. Haploid *S. cerevisiae* null mutants were constructed by replacing the genes coding for the enzymes malate synthase (*MLS1*), isocitrate lyase (*ICL1*), phosphoenolpyruvate carboxykinase (*PCK1*) and the fructose-1,6-bisphosphatase (*FBP1*) by the short flanking homology (SFH) method (Wach et al., 1994). SFH deletion cassettes were made by PCR with primers homologous to both the kanamycin deletion cassette of plasmid pUG6 (Güldener et al., 1996) and the gene(s) of interest as described recently (Luttik et al., 1998). SFH-PCR products were first transformed in the diploid strain CEN.PK122 and subsequently the corresponding haploid deletion strains were obtained by tetrad analysis. Mating type and the correct integration of the kanamycin deletion cassette were verified by PCR as described previously (Luttik et al., 1998). Standard techniques and media for genetic modification of *S. cerevisiae* were used (Ausubel et al., 1989). All strains designated as CEN are isogenic to the *S. cerevisiae* strain CEN.PK2 (Entian and Kötter, 1998; van Dijken et al., 2000).

<table>
<thead>
<tr>
<th>strain name</th>
<th>relevant genotype*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEN.PK113-7D</td>
<td>MATa URA3 HIS3 LEU2 TRP1 MAL2-8° SUC2</td>
</tr>
<tr>
<td>CEN.PK122</td>
<td>MATa/MATa URA3/URA3 HIS3/HIS3 LEU2/LEU2 TRP1/TRP1 MAL2-8° /MAL2-8° SUC2/SUC2</td>
</tr>
<tr>
<td>CEN.PK229-4D</td>
<td>MATa URA3 HIS3 LEU2 TRP1 MAL2-8° SUC2 icl1(41,1630)::loxP-Kan-loxP</td>
</tr>
<tr>
<td>CEN.PK233-1C</td>
<td>MATa URA3 HIS3 LEU2 TRP1 MAL2-8° SUC2 mls1(41,1610)::loxP-Kan-loxP</td>
</tr>
<tr>
<td>CEN.PK255-1D</td>
<td>MATa URA3 HIS3 LEU2 TRP1 MAL2-8° SUC2 fbp1(41,880)::loxP-Kan-loxP</td>
</tr>
<tr>
<td>CEN.PK487-1B</td>
<td>MATa URA3 HIS3 LEU2 TRP1 MAL2-8° SUC2 cat8(4,4299)::loxP-Kan-loxP</td>
</tr>
<tr>
<td>CEN.IS46-1A</td>
<td>MATa URA3 HIS3 LEU2 TRP1 MAL2-8° SUC2 ppc1(41,1610)::loxP-Kan-loxP</td>
</tr>
</tbody>
</table>

*The numbers in parentheses indicate the deleted nucleotides (ATG = 1) of the corresponding genes. All strains listed can be obtained from EUROSCARF (European Saccharomyces cerevisiae Archive for Functional Analysis) at http://www.rz.uni-frankfurt.de/FB/fb16/mikro/euroscarf/index.html.*

Table I: yeast strains used in this study
Strain cultivation and analysis

The haploid wild type strain *S. cerevisiae* CEN.PK113-7D was used as reference for the derived knockout mutants. The null mutants in the enzymes of the glyoxylate cycle enzymes malate synthase (mls) and isocitrate lyase (icl) and of the gluconeogenic enzymes phosphoenolpyruvate carboxykinase (pck) and fructose-1,6-bisphosphatase (fbp) were made via the kanamycin method (Güldener et al., 1996) in the diploid *S. cerevisiae* wild typ strain CEN.PK122. The following obtained haploid mutant strains of *S. cerevisiae* were used for the experiments: CEN.PK 233-1C (*mls1Δ*), CEN.PK 229-4D (*icl1Δ*), CEN.PK 46-1A (*pck1Δ*), CEN.PK 255-1D (*fbp1Δ*).

The wild type strain was grown aerobically in carbon-limited chemostat cultures using 2 litre Applikon fermentors at different dilution rates between 0.025 and 0.2 hr⁻¹, 30°C and pH 5.0 on defined mineral media according to Verduyn et al. (Verduyn et al., 1992) with either glucose, ethanol or acetate as carbon source. The mutants were grown under the same conditions on glucose and ethanol in 1.5 l working volume at a dilution rate of D = 0.1 hr⁻¹ using Braun fermentors (B.Braun biotech international, 2 l). Dissolved oxygen tension remained always above 30%. The fermentors were inoculated with approximately 50 ml preculture, grown overnight on the same mineral medium and 10 g/l glucose as carbon-source.

Experiments with varying glucose-ethanol feed were performed with a constant total carbon concentration of 250 mM. For the determination of the energetic parameters the feed concentration varied between 180 and 480 mM carbon for glucose, ethanol was fed at a concentration of 350 mM, acetate at 270 mM. Carbon balances were checked routinely and closed with a recovery of more than 95% for all cultures.

The fermentors were aerated with an airflow of approximately 1vvm. Off-gas analysis for oxygen and carbon dioxide was performed only for the wild typ cells using a Servomex 1100A Oxygen Analyser (Taylor Servomex Co., USA) and a Beckman 864 infrared detector (Rosemount Analytics, Santa Clara, CA, USA) respectively. Biomass concentrations were determined through filtration on nitro-cellulose filters (pore size 0.45μm, Gelman Sciences, Ann Arbor, MI, USA).

Biomass was harvested after reaching steady state. A culture was considered to be in steady state if the biomass concentration was constant for at least five dilution times and the biomass was devoid of cell cycle associated oscillations (Duboc et al., 1996). The cells were spun down and washed with demineralized water. The biomass was stored at -20°C until being freeze-dried for 48 hours. The biomass was then further dried at 70°C for 48 hours and stored at room temperature in a desiccator to obtain a reproducible reference state until further analysis.

Biomass was analysed for its elemental and constituent content as described earlier. Biomass was assumed to contain protein, carbohydrates, lipids, RNA, DNA, water, and
Characterisation of null mutants of glyoxylate cycle and gluconeogenic enzymes

metals. Results of the elemental and the constituent analysis were matched using statistical reconciliation (Lange and Heijnen, 2001).

Enzyme assays
For enzyme assays 5 ml culture were harvested after reaching steady state by centrifugation. The biomass was stored at -20°C until being used. Crude extracts were prepared using glass beads for breaking the cells as described by Ciriacy et al. (Ciriacy, 1973). The crude extract was then directly used for the enzyme assays.

Malate synthase and isocitrate lyase activity were determined according to Dixon and Kornberg (Dixon and Kornberg, 1959), phosphoenolpyruvate carboxykinase activity according to Hansen et al. (Hansen et al., 1976) and fructosebisphosphatase activity according to Gancedo and Gancedo (Gancedo and Gancedo, 1971). Specific activities were computed as mU/mg protein, determined according to the microbiuret method (Zamenhoff, 1957) using bovine serum albumin as a standard.

Metabolic network analysis
Metabolic network analysis was performed with a network based on van Gulik and Heijnen (van Gulik and Heijnen, 1995). Further reactions were included for the formation of the phosphatidates, precursors of phospholipids containing ethanolamine, choline, and serine. Different biomass-generating reactions were included to reflect the differences in measured biomass and protein composition in each steady state. Biomass, CO₂ and H₂O were assumed to be the only products of metabolism (Appendix).

Networks for glucose cultures were modelled using pyruvate carboxylase as anaplerotic reaction for the TCA cycle. NADPH was regenerated through the NADP dependent isocitrate dehydrogenase and the glucose-6-phosphate dehydrogenase with the pentose-phosphate pathway providing the required stoichiometric freedom for this cofactor. Networks for ethanol grown cultures included both NAD and NADP dependent aldehyde dehydrogenases. In this case the PP-pathway was restricted to provide pentose-phosphate precursors by removing the second transketolase, hereby removing a singularity introduced by the additional NAD dependant aldehyde dehydrogenase. Gluconeogenesis follows the known route with PEP carboxylase and fructose bisphosphatase, replacing the pyruvate carboxylase and the phosphofructose kinase. The reactions of the glyoxylate cycle isocitrate lyase and malate synthase fulfill the anaplerotic requirements for the growth on C₂-components. The network for acetate is similar to the one used for ethanol grown cultures with the sole exception of the use of the second transketolase to provide the freedom to generate the required NADPH in the PP-pathway.
In order to use the metabolic networks for predictions, the unknown energy requirements expressed in ATP of the assimilatory reactions and the amount of ATP spent on maintaining the current state of the cell have to be evaluated. These additional energetic requirements can be expressed with three energetic parameters: the P/O ratio, describing the efficiency of the oxidative phosphorylation, the growth related unknown ATP requirements for cell formation $k$, and the cell maintenance $m_{ATP}$ of the organism (van Gulik and Heijnen, 1995). This energy demand can be solved using the ATP balance for each metabolic network:

$$q_o \cdot \frac{P}{O} - \mu \cdot k - m_{ATP} = q_{ATP}^{net} \tag{1}$$

The specific rates $q$ per C-mol biomass can be easily determined through metabolic flux analysis with $q_o$ representing the moles of oxygen consumed for the reduction of NADH and FADH, and $q_{ATP}^{net}$ being the amount of known ATP used in the specified anabolic and catabolic reactions. The $q_{ATP}^{net}$ can be quantified by summarizing all reaction rates involving ATP except the regeneration through oxidative phosphorylation:

$$q_{ATP}^{net} = \sum v_{i,ATP} r_i \tag{2}$$

Each biological steady state thus generates its distinct equations (1) and (2). Using the appropriate metabolic network for equation (2), the energetic parameters P/O ratio, the growth related ATP requirements $k$, and the cell maintenance $m_{ATP}$ can be obtained from data of minimal three distinct fermentations of the wild type strain using linear regression techniques on equation (1). The accuracy of the estimated parameters is improved through the use of data from several chemostats with multiple substrates and a range of growth rates. The maintenance parameters $m_{ATP}$ and $k$ can be combined to $k'$ for modelling purposes at a constant growth rate.

$$k' = k + \frac{m_{ATP}}{\mu} \tag{3}$$

The flux analysis of the different null- mutants was done following the strategy outlined by van Gulik and Heijnen (van Gulik and Heijnen, 1995) and Vanrolleghem and Heijnen (Vanrolleghem and Heijnen, 1998) who successfully applied it for the sequential enzyme expression during growth of S. cerevisiae on mixtures of glucose and ethanol (de Jong-Gubbels et al., 1995). As proposed in these papers, a set of different metabolic networks was used to model the complete range of the glucose-ethanol ratio in the feed.
Results and Discussion

Biomass composition

Metabolic flux analysis requires the description of the biomass not only in terms of its elements, but also in terms of its molecular composition. Therefore, the average composition for *Saccharomyces cerevisiae* CEN.PK113-7D grown in carbon-limited chemostat at a dilution rate of $D = 0.1 \text{ hr}^{-1}$ was analysed for its elemental and molecular composition during growth on glucose, ethanol, and acetate as described previously (Lange and Heijnen, 2001). The elemental composition is presented in table II. The molecular composition is described in terms of the mass fractions $f$ of protein,

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$D$ ($\text{hr}^{-1}$)</th>
<th>$f_C$</th>
<th>$f_H$</th>
<th>$f_O$</th>
<th>$f_P$</th>
<th>$f_S$</th>
<th>$f_M$</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>glucose</td>
<td>0.105</td>
<td>45.5</td>
<td>6.7</td>
<td>7.9</td>
<td>36.2</td>
<td>1.03</td>
<td>0.23</td>
<td>2.52</td>
</tr>
<tr>
<td>ethanol</td>
<td>0.106</td>
<td>46.7</td>
<td>6.9</td>
<td>8.3</td>
<td>34.1</td>
<td>1.04</td>
<td>0.31</td>
<td>2.61</td>
</tr>
<tr>
<td>acetate</td>
<td>0.106</td>
<td>46.5</td>
<td>6.9</td>
<td>8.9</td>
<td>33.8</td>
<td>1.05</td>
<td>0.32</td>
<td>2.48</td>
</tr>
</tbody>
</table>

**Table II:** balanced elemental cell composition expressed as mass fractions $f$, in %

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$D$ ($\text{hr}^{-1}$)</th>
<th>$f_{\text{protein}}$</th>
<th>$f_{\text{carbohydr}}$</th>
<th>$f_{\text{lipids}}$</th>
<th>$f_{\text{RNA}}$</th>
<th>$f_{\text{DNA}}$</th>
<th>$f_{\text{Pi}}$</th>
<th>$f_{\text{SO4}}$</th>
<th>$f_{\text{water}}$</th>
<th>$f_{\text{metals}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>glucose</td>
<td>0.105</td>
<td>39.3</td>
<td>39.8</td>
<td>6.2</td>
<td>6.5</td>
<td>0.4</td>
<td>0.8</td>
<td>0.3</td>
<td>4.1</td>
<td>2.5</td>
</tr>
<tr>
<td>ethanol</td>
<td>0.106</td>
<td>42.2</td>
<td>33.7</td>
<td>9.8</td>
<td>6.1</td>
<td>0.5</td>
<td>0.7</td>
<td>0.5</td>
<td>4.0</td>
<td>2.6</td>
</tr>
<tr>
<td>acetate</td>
<td>0.106</td>
<td>45.1</td>
<td>30.9</td>
<td>8.7</td>
<td>6.8</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5</td>
<td>4.4</td>
<td>2.5</td>
</tr>
</tbody>
</table>

**Table III:** balanced elemental cell composition expressed as mass fractions $f$, in %

carbohydrates, lipids, RNA, DNA, phosphate, sulphate, metals, and water (table III). The different substrates lead to significant changes in the biomass composition even while maintaining a constant growth rate. Glucose grown cultures had a higher content of carbohydrates at $D = 0.1 \text{ hr}^{-1}$, while biomass grown on the C$_2$-components ethanol and acetate was richer in lipids and proteins. Details of the reconciled biomass composition are given in table III. Furthermore, the protein was analysed for its amino acid contents, an average composition (table IV) was used for the calculations since no significant differences were found between the cultures. Biomass composition also shifted e.g. to higher protein contents with increasing dilution rate as shown earlier (Lange and Heijnen, 2001), these changes were taken into account during the

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Mass %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>9.77</td>
</tr>
<tr>
<td>Arg</td>
<td>3.86</td>
</tr>
<tr>
<td>Asx</td>
<td>9.28</td>
</tr>
<tr>
<td>Cys</td>
<td>0.14</td>
</tr>
<tr>
<td>Glx</td>
<td>15.48</td>
</tr>
<tr>
<td>Gly</td>
<td>8.89</td>
</tr>
<tr>
<td>His</td>
<td>1.93</td>
</tr>
<tr>
<td>Ile</td>
<td>5.89</td>
</tr>
<tr>
<td>Leu</td>
<td>8.01</td>
</tr>
<tr>
<td>Lys</td>
<td>6.57</td>
</tr>
</tbody>
</table>

*Asx = Asp + Asn; Glx = Gln + Gln*

**Table IV:** average measured amino acid composition of protein in mol %
determination of the energetic parameters of the wild type. For the simulation of the mixed substrate growth, a linear interpolation of the values obtained for glucose and ethanol was used to reflecting the different biomass composition for pure glucose or ethanol growth respectively. The ratio between both was linearly linked to the glucose-ethanol ratio in the feed.

**Determination of energetic parameters**

The energetic parameters $P/O$ ratio, $k_m$, and $m_{ATP}$ of the strain were calculated based on results of 14 chemostats, 12 of which were grown on glucose under different growth rates and two with a growth rate of $D = 0.1\text{hr}^{-1}$ grown on ethanol and acetate respectively. The measurements underwent chi-square tests, the constraints being provided by the metabolic network. All data sets were consistent within a 95%-confidence interval. Statistical reconciliation of the measurements were done according to van der Heijden (van der Heijden et al., 1994) before the calculation of the energetic parameters. For $D = 0.1\text{hr}^{-1}$ the average biomass yield was determined as 0.600, 0.623, and 0.382 C-mol$_{biomass}$/C-mol$_{substrate}$ for glucose, ethanol, and acetate respectively. The yield on glucose increased from 0.483 to 0.624 C-mol$_{biomass}$/C-mol$_{substrate}$ with dilution rates increasing from 0.02 to 0.2hr$^{-1}$ (table V). Subsequently, the $P/O$ ratio was determined for the metabolic network presented in appendix A as 1.45, $k_m$ as 655 mmol$_{ATP}$/C-mol$_{biomass}$, and $m_{ATP}$ was evaluated to 56 mmol ATP/mol biomass/hr. These values were then used for modelling of the deletion mutants on the range of mixed ethanol/glucose feeds.

<table>
<thead>
<tr>
<th>substrate</th>
<th>dilution rate hr$^{-1}$</th>
<th>RQ mol/mol</th>
<th>$Y_{XS}$ C-mol/C-mol</th>
<th>$Y_{XO}$ C-mol/mol-O$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethanol</td>
<td>0.106</td>
<td>0.45</td>
<td>0.623</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.022</td>
<td>1.06</td>
<td>0.483</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>0.052</td>
<td>1.08</td>
<td>0.570</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>0.087</td>
<td>1.07</td>
<td>0.586</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td>0.099</td>
<td>1.07</td>
<td>0.591</td>
<td>1.55</td>
</tr>
<tr>
<td></td>
<td>0.104</td>
<td>1.08</td>
<td>0.620</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>0.107</td>
<td>1.07</td>
<td>0.598</td>
<td>1.59</td>
</tr>
<tr>
<td>glucose</td>
<td>0.107</td>
<td>1.07</td>
<td>0.613</td>
<td>1.70</td>
</tr>
<tr>
<td></td>
<td>0.107</td>
<td>1.07</td>
<td>0.591</td>
<td>1.54</td>
</tr>
<tr>
<td></td>
<td>0.108</td>
<td>1.06</td>
<td>0.571</td>
<td>1.41</td>
</tr>
<tr>
<td></td>
<td>0.126</td>
<td>1.08</td>
<td>0.608</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>0.158</td>
<td>1.09</td>
<td>0.620</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>0.211</td>
<td>1.09</td>
<td>0.624</td>
<td>1.80</td>
</tr>
</tbody>
</table>

Table V: reconciled respiratory quotient and yields on substrate and oxygen of *S. cerevisiae* CEN.PK113-7D

The determined $P/O$ ratio is substantially higher than the values previously reported for *S. cerevisiae* (Vanrolleghem et al., 1996; van Gulik and Heijnen, 1995). However, the set of energetic parameters is dependant on the chosen network and can thus not be easily
Characterisation of null mutants of glyoxylate cycle and gluconeogenic enzymes

compared or transferred to other networks. Further improvements we introduced as the inclusion of ion transport, and a larger range of dilution rates had only minor effects on the energetic parameters. Furthermore, the biomass composition was measured for each steady state, reconciled values according to Lange and Heijnen (Lange and Heijnen, 2001) were used in the presented model. Calculations based on a single average biomass composition resulted in a 4% lower P/O value.

However, since the purpose of the used model is not to obtain a perfect description of the cell’s energetics, but to provide an easily applicable tool, we tested the influence of the different parameters. Since a lower P/O ratio is compensated for by the reduced requirements for maintenance, we found only small variations of less than 1.5% during the switching point determination when using a P/O value of 1.2.

Modelling of growth on glucose-ethanol mixtures

Five distinct networks were used to model the growth of the deletion mutants on glucose-ethanol mixtures, using a similar approach as suggested by van Gulik and Heijnen (van Gulik and Heijnen, 1995). Each network spans a distinct range of glucose:ethanol ratios in the feed and is limited at each side by a certain ratio, the switching point, which requires a different set of enzymes to completely metabolise both substrates. With increasing ethanol content in the feed, the cell requires first simultaneously the two additional enzymes of the glyoxylate cycle, isocitrate lyase and malate synthase, followed by the gluconeogenic enzymes PEP carboxylase and fructose-1,6-bisphosphatase. The three switching points were determined mathematically by solving an under-determined network for a ‘zero’ flux for the particular reaction. For the modelling of the mutants, we assumed the same metabolism as for the wild type as long as the ratio of ethanol to glucose remained below the switching point of the missing enzyme. For higher ratios, the glucose will continue to be completely metabolised, but only a fraction of the provided ethanol will be taken up by the cell, thus effectively fixing the uptake ratio of glucose to ethanol to the value of the switching point. This will subsequently lead to a decrease of the apparent yield with higher ethanol fraction and to an accumulation of ethanol in the medium.

Mutants lacking either the malate synthase or the isocitrate lyase can use ethanol only to provide acetyl-CoA but not for anaplerosis. This has to be provided for by the glucose metabolism. Under the conditions used, i.e., a growth rate of 0.1 h⁻¹ the metabolic network model predicts a maximum ethanol:glucose ratio of 48:52 on C-mol basis. Further ethanol will not be metabolised by the mls1Δ or icl1Δ mutants and will therefore remain in the broth. A strain lacking PEP carboxykinase can replenish the TCA-cycle from ethanol, but requires alternative substrates for glycolytic precursors. From this, the maximum uptake ratio of ethanol to total carbon for the pck1Δ mutant was calculated as 59 % (Cmol ethanol/Cmol carbon). The fructose bisphosphatase null mutant requires glucose for its C₆-precursors, the model predicts here a maximum ethanol fraction of 74 %.

While the energetic parameters are very sensitive to the chosen metabolic network, the switching points remain almost the same, network variations rarely changed the here presented points by more than 1%.
Verification of the models by chemostat cultivation

The results of the metabolic flux model were compared with the actual measurements of the yield and the residual ethanol content in chemostat cultures. For the wild type the model anticipated an approximately constant biomass yield and no residual ethanol for the whole range of different ethanol to glucose ratios (figures 1A and 2A). For the different mutants, the model predicted a decrease in the biomass yield accompanied by the accumulation of ethanol for ratios exceeding the maximum ratio for the particular mutant. This was indeed observed in the chemostat experiments with the mutants. The icl1Δ and the mls1Δ mutants both exhibited the same linear decrease of the biomass concentration for ethanol fractions higher than 50% until complete disappearance for a 100% ethanol feed, agreeing well with the model prediction of 47% (figure 1B). The increase in the residual ethanol exhibited a slight lag compared to the predicted values. The lag might simply be attributed to evaporation. However, even at an ethanol content in the feed of 90% the predicted residual ethanol concentration (202 CmM) is almost 20% higher than the measured one (170 CmM)(figure 2B).

**Figure 1:** predicted and measured biomass yields for *S. cerevisiae* grown in carbon-limited chemostat over the ratio of ethanol to glucose in the feed. A.: CEN.PK113-7D (wild type); B.: CEN.PK 233-1C (mls1Δ) and CEN.PK 229-4D (icl1Δ); C.: CEN.PK 46-1A (pck1Δ); D.: CEN.PK 255-1D (fpb1Δ)

For the PEP carboxykinase deletion mutant, the ethanol again starts to accumulate slightly later than predicted, but the increase follows the same slope as predicted. The measured
biomass concentration starts to drop significantly below that of the wild type when the ethanol fraction in the feed exceeds 60%. The model predicts a sharp linear decrease at an ethanol fraction above 59%, thus model and measurements agree very well (figure 1C).

The fructose bisphosphatase deletion mutant shows a continuous slow drop in biomass concentration already earlier than predicted by the model, which forecasted a sharp decline at an ethanol concentration above 74%. The measured residual ethanol concentration of 29.1CmM (0.67 g/l) for the mutant at an ethanol concentration in the feed of 90% is significantly lower than the predicted value of 141 CmM (3.26 g/l), indicating a discrepancy between model prediction and real life behaviour of this mutant (figures 1D and 2D). Obviously, the cell's metabolism reacts in a more gradual fashion than envisioned by the model, possibly caused by changes in the biomass composition.

![Graph A](image1.png)  ![Graph B](image2.png)  ![Graph C](image3.png)  ![Graph D](image4.png)

**Figure 2**: predicted and measured residual ethanol for *S. cerevisiae* grown in carbon-limited chemostat over the ratio of ethanol to glucose in the feed. A.: CEN.PK113-7D (wild type); B.: CEN.PK 233-1C (*msl1Δ*) and CEN.PK 229-4D (*icl1Δ*); C.: CEN.PK 46-1A (*pck1Δ*); D.: CEN.PK 255-1D (*fhp1Δ*)

In figure 3 the measured and predicted ethanol fraction of the consumed substrate has been plotted against the supplied ethanol fraction in the feed. All ethanol is consumed until the supplied ethanol fraction reaches the switching point. For higher ethanol fractions in the feed, the model predicts a constant consumed ethanol fraction. For the *fhp1Δ* mutant the measurements and predictions do not compare very well. The measured data for the *pck1Δ* and *icl1Δ/msl1Δ* mutants show some scatter. The average ethanol fraction of the consumed substrate is in both cases somewhat higher than predicted by the model.

55
Both the delayed onset of the ethanol accumulation as well as the slightly higher consumed ethanol fraction as observed in all cultures compared to the predicted value might be simply due to a shift in the biomass composition of the model. The yeast cell has the ability to adapt its composition according to its substrate source as could be seen from the different biomass composition for glucose, ethanol, and acetate. It is therefore easily conceivable that a mutant unable to provide the required precursors for the anabolism of carbohydrates will instead have a relative higher abundance in other storage polymers e.g. lipids, and will thus be able to utilize a larger ethanol fraction.

![Graph showing ethanol consumption vs. supplied carbon](image)

**Figure 3:** consumed versus provided ethanol in the feed for *S. cerevisiae* in carbon-limited chemostat. ★ CEN.PK113-7D (wild type), ◆ CEN.PK233-1C (mls1Δ) and CEN.PK229-4D (icl1Δ), ● CEN.PK46-1A (pck1Δ), ▲ CEN.PK255-1D (fbp1Δ)

**Conclusions**

A metabolic model based on four distinct networks was able to predict the quantitative growth of the icl1Δ, the mls1Δ and the pck1Δ mutants on mixtures of glucose and ethanol correctly. The model was based solely on the known pathways in yeast and the energetic parameters obtained in chemostat cultivations of the corresponding wild type strain of *Saccharomyces cerevisiae* CEN.PK113-7D. Quantitative agreement was less accurate for the fbp1Δ mutant, but correct in its tendency. This demonstrates the capacity of metabolic flux analysis to correctly predict proposed genetical changes *in silico* before these are realized *in vivo*.

The experiments show the specific glucose-ethanol concentration from which the activity of the single enzyme is needed to metabolize both glucose and ethanol completely. These results also verify the statement from former publications (de Jong-Gubbels et al., 1995) that the activity of the four enzymes is needed one after the other beginning with isocitrate
lyase and malate synthase over phosphoenolpyruvate carboxykinase to fructose-1,6-bisphosphatase with rising ethanol content in mixed media.

Regulation of the enzymatic activity in vivo seems mainly to be caused by post-translational regulatory mechanisms such as allosteric effects and phosphorylation of the proteins, since western blot experiments with the mentioned proteins (data not shown) reveals comparable protein concentrations of all four genes over the experiments. Experiments with a deletion strain in the gene for Cat8p, an activator for the gluconeogenic genes (Hedges et al., 1995), shows behaviour like the deletion strains for isocitrate lyase and malate synthase (data not shown). This result gives evidence that gene regulation is the same on the transcriptional level for all four genes. This conclusion is further supported by the fact, that the promotor sections of the four genes have the same regulatory elements (Proft et al., 1995; Hedges et al., 1995; Schüller, 1994; Caspary et al., 1997).

The results point to a tight post-translational regulation of the gluconeogenic enzymes which are expressed during growth on a glucose-ethanol mixture. The regulation of their in-vivo activity is such as to prevent futile cycling. The control seem to rest solely with the available glucose and is independent of the external ethanol concentration.

List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$</td>
<td>dilution rate</td>
<td>hr$^{-1}$</td>
</tr>
<tr>
<td>$f_i$</td>
<td>mass fraction of component $i$ in biomass</td>
<td>%</td>
</tr>
<tr>
<td>$k'$</td>
<td>combined maintenance coefficient</td>
<td>mol ATP/C-mol biomass</td>
</tr>
<tr>
<td>$k_m$</td>
<td>growth related maintenance</td>
<td>mol ATP/C-mol biomass</td>
</tr>
<tr>
<td>$m_{ATP}$</td>
<td>maintenance expressed in mol ATP</td>
<td>mol ATP/C-mol biomass/hr</td>
</tr>
<tr>
<td>$P/O$</td>
<td>yield of ATP per oxygen atom</td>
<td>-</td>
</tr>
<tr>
<td>$q_i$</td>
<td>specific flow rate of component $i$</td>
<td>mol/C-mol biomass/hr</td>
</tr>
<tr>
<td>$r_i$</td>
<td>specific reaction rate $i$</td>
<td>mol/C-mol biomass/hr</td>
</tr>
<tr>
<td>$\mu$</td>
<td>growth rate</td>
<td>hr$^{-1}$</td>
</tr>
<tr>
<td>$v_{ij}$</td>
<td>stoichiometric coefficient of comp. $j$ in reaction $i$</td>
<td>-</td>
</tr>
</tbody>
</table>

superscripts

- **net**: related to the metabolic network
- **cat**: related to catabolism
- **o**: oxygen
Appendix: list of metabolic reactions

**amino acid synthesis**

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Formulas</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALA N-trans</td>
<td>PYR + GLM -&gt; OGL + ALA</td>
</tr>
<tr>
<td>ARG syn</td>
<td>ORN + CARP + ATP + ASP -&gt; AMP + ARG + FUM + 3 H + Pi + PPI</td>
</tr>
<tr>
<td>ASN syn</td>
<td>H2O + GLN + ATP + ASP -&gt; ADP + ASN + GLM + H + Pi</td>
</tr>
<tr>
<td>ASP N-trans</td>
<td>OXACT + GLM -&gt; OGL + ASP</td>
</tr>
<tr>
<td>ASP kin</td>
<td>2 NADPH + 2 H + ATP + ASP -&gt; ADP + HSER + 2 NADP + Pi</td>
</tr>
<tr>
<td>CARP syn</td>
<td>2 H2O + GLN + CO2 + 2 ATP -&gt; 2 ADP + CARP + GLM + 3 H + Pi</td>
</tr>
<tr>
<td>CYS syn</td>
<td>HCYS + SER + 2 H -&gt; CYS + NH4 + OBU</td>
</tr>
<tr>
<td>GLM deh</td>
<td>NH4 + NADPH + H + OGL -&gt; GLM + H2O + NADP</td>
</tr>
<tr>
<td>GLM N-lig</td>
<td>NH4 + GLM + ATP -&gt; ADP + GLN + H + Pi</td>
</tr>
<tr>
<td>GLY transf</td>
<td>THF + SER -&gt; GLY + H2O + METHF</td>
</tr>
<tr>
<td>HCYS syn</td>
<td>HCYS + HSER + ACCoA -&gt; ACT + CoA + 2 H + HCYS</td>
</tr>
<tr>
<td>HIS syn</td>
<td>FRPP + 2 NAD + 3 H2O + GLN + ATP -&gt; OGL + 6 H + HIS + 2 NADH + Pi + 2 PPI + AICAR</td>
</tr>
<tr>
<td>ILE N-trans</td>
<td>OBU + PYR + NADPH + H + GLM -&gt; OGL + CO2 + H2O + ILE + NADP</td>
</tr>
<tr>
<td>LEU N-trans</td>
<td>NAD + H2O + GLM + OIV + ACCoA -&gt; OGL + CO2 + CoA + H + LEU + NADH</td>
</tr>
<tr>
<td>LYS syn</td>
<td>2 NADPH + 2 NAD + H2O + 2 GLM + ATP + ACCoA -&gt; OGL + AMP + CO2 + CoA + H + LYS + 2 NADH + 2 NADP + PPI</td>
</tr>
<tr>
<td>MET syn</td>
<td>HCYS + MYTHF + H -&gt; MET + THF</td>
</tr>
<tr>
<td>OIV syn</td>
<td>2 PYR + NADPH + 2 H -&gt; OIV + CO2 + H2O + NADP</td>
</tr>
<tr>
<td>ORN syn</td>
<td>NADPH + H2O + 2 GLM + ATP + ACCoA -&gt; ACT + ADP + OGL + CoA + NADP + Pi + ORN</td>
</tr>
<tr>
<td>PHE syn</td>
<td>H + GLM + CHO -&gt; OGL + CO2 + H2O + PHE</td>
</tr>
<tr>
<td>PRO deh</td>
<td>2 NADPH + 2 H + GLM + ATP -&gt; ADP + H2O + 2 NADP + Pi + PRO</td>
</tr>
<tr>
<td>SER syn</td>
<td>3PG + NAD + H2O + GLM -&gt; OGL + H + NADH + Pi + SER</td>
</tr>
<tr>
<td>SHI path</td>
<td>2 PEP + NADPH + E4P + ATP -&gt; ADP + CHO + NADP + 4 Pi</td>
</tr>
<tr>
<td>THR ald</td>
<td>THR -&gt; ACTAL + GLY</td>
</tr>
<tr>
<td>THR deh</td>
<td>THR + H -&gt; NH4 + OBU</td>
</tr>
<tr>
<td>THR syn</td>
<td>HSER + H2O + ATP -&gt; ADP + H + Pi + THR</td>
</tr>
<tr>
<td>TRP syn</td>
<td>SER + PRPP + GLN + CHO -&gt; CO2 + GAP + GLM + H + H2O + 2 Pi + PYR + TRP</td>
</tr>
<tr>
<td>TYR syn</td>
<td>NADP + GLM + CHO -&gt; OGL + CO2 + NADPH + TYR</td>
</tr>
<tr>
<td>VAL N-trans</td>
<td>GLM + OIV -&gt; OGL + VAL</td>
</tr>
</tbody>
</table>

**biomass formation**

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Formulas</th>
</tr>
</thead>
<tbody>
<tr>
<td>bio-A1</td>
<td>0.1716 metal + 0.0385 DNA + 0.13456 LIPID + 0.00188 SO4 + 0.05185 RNA + 0.294537 CARBYHD + 0.51428 PROT + 0.0163 Pi + 0.0329 H2O -&gt; biomass-A1s</td>
</tr>
<tr>
<td>biom-E1</td>
<td>0.01799 metal + 0.00375 DNA + 0.15073 LIPID + 0.00186 SO4 + 0.04611 RNA + 0.32105 CARBYHD + 0.47917 PROT + 0.00196 Pi + 0.05674 H2O -&gt; biomass-E1s</td>
</tr>
<tr>
<td>biom-G1</td>
<td>0.01812 metal + 0.00354 DNA + 0.12682 LIPID + 0.00114 SO4 + 0.03005 RNA + 0.43765 CARBYHD + 0.40186 PROT + 0.00376 Pi + 0.05998 H2O -&gt; biomass-G1s</td>
</tr>
<tr>
<td>biom-G2</td>
<td>0.01759 metal + 0.0035 DNA + 0.12144 LIPID + 0.00101 SO4 + 0.03521 RNA + 0.39508 CARBYHD + 0.44469 PROT + 0.0032 Pi + 0.05661 H2O -&gt; biomass-G2s</td>
</tr>
<tr>
<td>biom-G3</td>
<td>0.01825 metal + 0.00353 DNA + 0.09754 LIPID + 0.00153 SO4 + 0.04175 RNA + 0.39545 CARBYHD + 0.46165 PROT + 0.00295 Pi + 0.0545 H2O -&gt; biomass-G3s</td>
</tr>
<tr>
<td>biom-G4</td>
<td>0.01797 metal + 0.00362 DNA + 0.10086 LIPID + 0.00094 SO4 + 0.04701 RNA + 0.39766 CARBYHD + 0.45077 PROT + 0.00245 Pi + 0.05611 H2O -&gt; biomass-G4s</td>
</tr>
<tr>
<td>biom-G5</td>
<td>0.01813 metal + 0.00404 DNA + 0.10711 LIPID + 0.00186 SO4 + 0.04806 RNA + 0.38846 CARBYHD + 0.45225 PROT + 0.00227 Pi + 0.0595 H2O -&gt; biomass-G5s</td>
</tr>
<tr>
<td>biom-G6</td>
<td>0.01771 metal + 0.00352 DNA + 0.09751 LIPID + 0.00092 SO4 + 0.05334 RNA + 0.3772 CARBYHD + 0.46835 PROT + 0.05874 H2O -&gt; 0.00197 Pi + biomass-G6s</td>
</tr>
<tr>
<td>biom-G7</td>
<td>0.01775 metal + 0.00355 DNA + 0.09786 LIPID + 0.00097 SO4 + 0.04855 RNA + 0.4046 CARBYHD + 0.44537 PROT + 0.00243 Pi + 0.06556 H2O -&gt; biomass-G7s</td>
</tr>
<tr>
<td>biom-G8</td>
<td>0.01744 metal + 0.00291 DNA + 0.09669 LIPID + 0.00087 SO4 + 0.0514 RNA + 0.37653 CARBYHD + 0.47239 PROT + 0.0021 Pi + 0.05991 H2O -&gt; biomass-G8s</td>
</tr>
</tbody>
</table>

58
Characterisation of null mutants of glyoxylate cycle and gluconeogenic enzymes

**biom-G9**
- .01825 metal + .00376 DNA + .09084 LIPID + .00094 SO4 + .05429 RNA + .38999
- CARBHYD + .46104 PROT + .00198 Pi + .06367 H2O -> biom-G9

**biom-G10**
- .0172 metal + .00351 DNA + .1092 LIPID + .00065 SO4 + .04448 RNA + .377
- CARBHYD + .46573 PROT + .00262 Pi + .06395 H2O -> biom-G10

**biom-G11**
- .01764 metal + .00352 DNA + .11022 LIPID + .00154 SO4 + .0461 RNA + .3492
- CARBHYD + .49088 PROT + .00258 Pi + .06973 H2O -> biom-G11

**biom-G12**
- .01748 metal + .00351 DNA + .10468 LIPID + .00179 SO4 + .05141 RNA + .33004
- CARBHYD + .50728 PROT + .00203 Pi + .0637 H2O -> biom-G12

**C-1 metabolism**
- DHF red: DHF + NADPH + H -> NADP + THF
- MTHF deh: NADPH + H + FTHF -> H2O + MTHF + NADP
- MTHF red: NADPH + MTHF + H -> MYTHF + NADP

**Catabolism**
- SO4 ass: SO4 + 4 NADPH + 3 H + 2 ATP -> ADP + AMP + 2 H2O + 4 NADP + Pi + Pi
  + H2S

**Gluconeogenesis**
- F16P phos: F16P + H2O -> F6P + Pi
- ICIT lyu: ICIT -> GLYO + SUC
- MAL syn: H2O + GLYO + ACCoA -> CoA + H + MAL
- PEP ck: OXACT + ATP -> ADP + CO2 + PEP

**Glycolysis, lower**
- Enol: 2PG -> H2O + PEP
- GAP deh: Pi + NAD + GAP -> H + NADH + 13PG
- 13PG kin: 13PG + ADP -> ATP + 3PG
- 3PG mut: 3PG -> 2PG
- PYR kin: PEP + H + ADP -> ATP + PYR

**Glycolysis, upper**
- F16P ald: F16P -> GAP + DHAP
- G6P iso: G6P -> F6P
- HX kin: GLUC + ATP -> ADP + G6P + H
- PF kin: F6P + ATP -> ADP + H + F16P
- TP iso: DHAP -> GAP

**Lipid synthesis**
- ACCoA carb: H2O + CO2 + ATP + ACCoA -> ADP + 2 H + Pi + MACoA
- avFA form: 4.4 PLLM-CoA + 1.4 PLLM-CoA + 1.7 OLE-CoA + STE-CoA -> 8.5 avFA-CoA
- avPL form: 4 PHD-ETA + 11 PHD-CHO + 3 PHD-SER -> 18 avPL
- FAT form: avFA-CoA + PHD + H2O -> CoA + Pi + FAT
- GOH3P trans: 2 avFA-CoA + GOH3P -> 2 CoA + PHD
- GOH deh: DHAP + NADH + H -> NAD + GOH3P
- MET Atrans: MET + 2 H2O + ATP -> H + 3 Pi + SAM
- PLM desat: PLM-CoA + O2 + NADPH + H -> 2 H2O + NADP + PLM-CoA
- PLM lig: PLM + H2O + CoA + ATP -> AMP + H + 2 Pi + PLM-CoA
- PLM syn: 7 MACoA + 14 NADPH + 20 H + ACCoA -> 7 CO2 + 8 CoA + 6 H2O + 14 NADP + PLM
- PHD-C trans: PHD + H2O + CTP -> 2 Pi + CMP-DGOH
- PHD-EA meth: 3 SAM + PHD-ETA -> 3 H + PHD-CHO + 3 SAH
- PHD-SER dec: PHD-SER -> CO2 + PHD-ETA
- PHD-SER syn: CMP-DGOH + SER -> CMP + PHD-SER
- SAH hyd: SAH + H2O -> H + HCYS + A
- STE desat: STE-CoA + O2 + NADPH + H -> 2 H2O + NADP + OLE-CoA
- STE ligase: STE + H2O + CoA + ATP -> AMP + H + 2 Pi + STE-CoA
- STE syn: 8 MACoA + 16 NADPH + 23 H + ACCoA -> 8 CO2 + 9 CoA + 7 H2O + 16 NADP + STE

**Macromolecule synthesis**
- avAA form: .02382 ORN + .73257 VAL + .19598 TYR + .06492 TRP + .55657 THR + .5328 SER
  + .42243 PRO + .37583 PHE + .11378 MET + .65699 LYS + .8008 LEU + .58939 ILE

59
Chapter 3

\[ +0.19258 \text{ HIS} + 0.8918 \text{ GLY} + 1.0214 \text{ GLM} + 0.52618 \text{ GLN} + 0.01387 \text{ CYS} + 0.5199 \text{ ASP} \\
+ 0.40849 \text{ ASN} + 0.38557 \text{ ARG} + 0.97695 \text{ ALA} \rightarrow \text{ CARBH} \text{YD SYN} + \text{ H}_2 \text{O} + 6 \text{ G6P} + \text{ ATP} \\
\rightarrow \text{ ADP} + \text{ H} + 2 \text{ Pi} + 6 \text{ CARBH} \text{YD} \\
\]

DNA poly 
\[0.2 \text{ GTP} + 0.3 \text{ UTP} + 0.3 \text{ ADP} + 0.3 \text{ DHF} + 0.3 \text{ ATP} \rightarrow \text{ H}_2 \text{O} + \text{ NADP} + 9.8 \text{ DNA} + \text{ PPI} + 0.3 \text{ DHF} \\
\]

LIPID form 
\[0.45 \text{ avPL} + 0.55 \text{ FAT} \rightarrow 47.22 \text{ LIPID} \\
\]

PROT poly 
\[2 \text{ H}_2 \text{O} + 3 \text{ ATP} + 0.2 \text{ AA} + 0.3 \text{ ATP} \rightarrow 2 \text{ ADP} + \text{ AMP} + 0.4 \text{ H} + 0.2 \text{ Pi} + 4.81023 \text{ PROT} + \text{ PPI} \\
\]

RNA syn 
\[0.2 \text{ GTP} + 0.3 \text{ UTP} + 0.3 \text{ ATP} \rightarrow 9.5 \text{ RNA} + \text{ PPI} \\
\]

nucleotide synthesis

A kin 
\[\text{ H} + \text{ AMP} + \text{ ADP} \rightarrow \text{ ATP} + \text{ A} \\
\]

AMP kin 
\[\text{ ATP} + \text{ AMP} \rightarrow 2 \text{ ADP} \\
\]

AMP syn 
\[\text{ IMP} + \text{ ATP} + \text{ ASP} \rightarrow \text{ ADP} + \text{ AMP} + \text{ FUM} + 2 \text{ H} + 2 \text{ Pi} \\
\]

CTP syn 
\[\text{ UTP} + \text{ H}_2 \text{O} + \text{ GLN} + 2 \text{ ATP} \rightarrow \text{ ADP} + \text{ CTP} + \text{ GLM} + 2 \text{ H} + 2 \text{ Pi} \\
\]

CMP kin 
\[\text{ CMP} + \text{ ATP} \rightarrow \text{ ADP} + \text{ CDP} \\
\]

GMP syn 
\[\text{ NAD} + \text{ IMP} + 2 \text{ H}_2 \text{O} + \text{ GLN} + 2 \text{ ATP} \rightarrow \text{ AMP} + \text{ GLM} + \text{ GMP} + 4 \text{ H} + \text{ NADH} + \text{ PPI} \\
\]

GMP kin 
\[\text{ GMP} + \text{ ATP} \rightarrow \text{ ADP} + \text{ GDP} \\
\]

IMP syn 
\[\text{ AICAR} + \text{ FTHF} \rightarrow \text{ H}_2 \text{O} + \text{ IMP} + \text{ THF} \\
\]

GDP kin 
\[\text{ GDP} + \text{ ATP} \rightarrow \text{ ADP} + \text{ GTP} \\
\]

UDP kin 
\[\text{ UDP} + \text{ ATP} \rightarrow \text{ ADP} + \text{ UTP} \\
\]

CDP kin 
\[\text{ CDP} + \text{ ATP} \rightarrow \text{ ADP} + \text{ CTP} \\
\]

PRPP syn 
\[\text{ RIBU5P} + \text{ ATP} \rightarrow \text{ AMP} + \text{ H} + \text{ PRPP} \\
\]

AICAR syn 
\[\text{ PRPP} + 2 \text{ H}_2 \text{O} + \text{ GLY} + 2 \text{ GLN} + \text{ FTHF} + \text{ CO2} + 4 \text{ ATP} + \text{ ASP} \rightarrow 4 \text{ ADP} + \text{ FUM} + 2 \text{ GLM} + 8 \text{ H} + 4 \text{ PI} + \text{ THF} + \text{ PPI} + \text{ AICAR} \\
\]

PPI asc 
\[\text{ PPI} + \text{ H}_2 \text{O} \rightarrow 2 \text{ Pi} \\
\]

UMP syn 
\[\text{ PRPP} + 0.5 \text{ O2} + \text{ CARP} + \text{ ASP} \rightarrow \text{ CO2} + 2 \text{ H}_2 \text{O} + \text{ Pi} + \text{ UMP} + \text{ PPI} \\
\]

UMP kin 
\[\text{ UMP} + \text{ ATP} \rightarrow \text{ ADP} + \text{ UDP} \\
\]

oxidative phosphorylation

FAD reg 
\[0.5 \text{ O2} + \text{ FADH2} \rightarrow \text{ FAD} + \text{ H}_2 \text{O} \\
\]

NAD reg 
\[0.5 \text{ O2} + \text{ NADH} + 2 \text{ H} \rightarrow 2 \text{ H}_2 \text{O} + \text{ NAD} \\
\]

OX NADH 
\[1.46 \text{ Pi} + 0.5 \text{ O2} + \text{ NADH} + 2.46 \text{ H} + 1.46 \text{ ADP} \rightarrow 1.46 \text{ ATP} + 2.46 \text{ H}_2 \text{O} + \text{ NAD} \\
\]

OX FADH2 
\[1.46 \text{ Pi} + 0.5 \text{ O2} + 1.46 \text{ H} + \text{ FADH2} + 1.46 \text{ ADP} \rightarrow 1.46 \text{ ATP} + \text{ FAD} + 2.46 \text{ H}_2 \text{O} \\
\]

maintenance

\[\text{ H}_2 \text{O} + \text{ ATP} \rightarrow \text{ ADP} + \text{ H} + \text{ Pi} \]

pentose phosphate pathway

G6P deh 
\[2 \text{ NADP} + \text{ H}_2 \text{O} + 6 \text{ G6P} \rightarrow 2 \text{ CO2} + 2 \text{ H} + 2 \text{ NADPH} + 6 \text{ RIBU5P} \\
\]

RIBU iso 
\[\text{ RIBU5P} \rightarrow \text{ RIB5P} \\
\]

RIBUP epi 
\[\text{ RIBU5P} \rightarrow \text{ XYL5P} \\
\]

TA1 
\[\text{ SED7P} + \text{ GAP} \rightarrow \text{ E4P} + \text{ F6P} \\
\]

TK1 
\[\text{ XYL5P} + \text{ RIB5P} \rightarrow \text{ GAP} + \text{ SED7P} \\
\]

TK2 
\[\text{ XYL5P} + \text{ E4P} \rightarrow \text{ F6P} + \text{ GAP} \\
\]

pyruvate branchpoint

ACTAL deh_NAD 
\[\text{ NAD} + \text{ H}_2 \text{O} + \text{ ACTAL} \rightarrow \text{ ACT} + 2 \text{ H} + \text{ NADH} \\
\]

ACTAL deh_NADP 
\[\text{ NADP} + \text{ H}_2 \text{O} + \text{ ACTAL} \rightarrow \text{ ACT} + 2 \text{ H} + \text{ NADPH} \\
\]

ACCOA syn 
\[\text{ H}_2 \text{O} + \text{ CoA} + \text{ ATP} + \text{ ACT} \rightarrow \text{ ACCOA} + \text{ AMP} + \text{ H} + 2 \text{ Pi} \\
\]

ETOH deh 
\[\text{ NADH} + \text{ H} + \text{ ACTAL} \rightarrow \text{ ETOH} + \text{ NAD} \\
\]

LAC deh 
\[\text{ NAD} + \text{ LAC} \rightarrow \text{ H} + \text{ NADH} + \text{ PYR} \\
\]

PYR carb 
\[\text{ PYR} + \text{ H}_2 \text{O} + \text{ CO2} + \text{ ATP} \rightarrow \text{ ADP} + 2 \text{ H} + \text{ Pi} + \text{ OXACT} \\
\]

PYR dec 
\[\text{ PYR} + \text{ H} \rightarrow \text{ ACTAL} + \text{ CO2} \\
\]

PYR deh 
\[\text{ PYR} + \text{ NAD} + \text{ CoA} \rightarrow \text{ ACCOA} + \text{ CO2} + \text{ NADH} \\
\]

TCA cycle

ACON 1 
\[\text{ CIT} \rightarrow \text{ H}_2 \text{O} + \text{ ACO} \\
\]

ACON 2 
\[\text{ ACO} + \text{ H}_2 \text{O} \rightarrow \text{ ICIT} \\
\]

CIT syn 
\[\text{ OXACT} + \text{ H}_2 \text{O} + \text{ ACCOA} \rightarrow \text{ CIT} + \text{ CoA} + \text{ H} \\
\]

FUM hy 
\[\text{ H}_2 \text{O} + \text{ FUM} \rightarrow \text{ MAL} \\
\]

ICIT deh_NAD 
\[\text{ NAD} + \text{ ICIT} \rightarrow \text{ OGL} + \text{ CO2} + \text{ NADH} \\
\]

ICIT deh_NADP 
\[\text{ NADP} + \text{ ICIT} \rightarrow \text{ OGL} + \text{ CO2} + \text{ NADPH} \\
\]

60
Characterisation of null mutants of glyoxylate cycle and gluconeogenic enzymes

MAL deh  NAD + MAL  $\rightarrow$  H + NADH + OXACT
OGL deh  NAD + CoA + OGL  $\rightarrow$  CO2 + NADH + SUCCOA
SUC deh  SUCCOA + Pi + ADP  $\rightarrow$  ATP + CoA + SUC

transport
ACT trans  2 H2 + ACT  $\rightarrow$  ACT + 2 H
ATPase  H2O + ATP  $\rightarrow$  ADP + H2 + Pi
ETOH trans  ETOH  $\rightarrow$  ETOH
CO2 diff  CO2  $\rightarrow$  CO2
GLUC fdiff  GLUC  $\rightarrow$  GLUC
GOH trans  H + GOH  $\rightarrow$  GOH + H2
H2O diff  H2O  $\rightarrow$  H2O
LAC trans  LAC + H2  $\rightarrow$  H + LAC
MET imp  metal  $\rightarrow$  metal
NH4 trans  NH4 + H2  $\rightarrow$  H + NH4
O2 diff  O2  $\rightarrow$  O2
Pi trans  Pi + 2 H2  $\rightarrow$  2 H + Pi
PYR trans  PYR + H2  $\rightarrow$  H + PYR
SO4 trans  SO4 + 3 H2  $\rightarrow$  3 H + SO4

References


Chapter 4

Setup of a compartmented metabolic network model
Setup of a compartmented metabolic network model

Abstract

In *Saccharomyces cerevisiae*, the metabolism is distributed over several organelles. This review discusses the localisation of the central pathways of anabolism and catabolism and attempts to connect these with the appropriate transport systems, giving an overview of the identified carriers in baker's yeast. Special emphasis is given to the reactions at the interface of glycolysis and TCA cycle, where several possibilities for the formation of acetyl-CoA are probed. The contribution of the different compartments during growth on glucose, ethanol, and acetate is investigated, the peroxisomes are found to be of minor importance for these substrates. Furthermore, the energetics of aerobic growth are briefly surveyed and its consequences for the metabolism outlined. Finally, a simple, compartmented metabolic network is presented, based on the current literature, which enables the calculations of metabolic fluxes across the membranes using a minimal set of the most likely carriers.
Introduction

In eukaryotic cells most enzymes are specifically located in either the cytosol or in one of the organelles. While the complete genome of the yeast *Saccharomyces cerevisiae* has been sequenced and about 4000 of the found 6300 open reading frames (ORFs) have been assigned an enzymatic, regulatory, or structural function, the localization of identified proteins is subject of ongoing research, knowledge is still incomplete and seemingly contradictory findings have been published for many proteins. Even less is known about the intracellular transport processes, which link together the pathways in different compartments.

Different mechanisms have been found to be responsible for the intracellular targeting of the active protein. For instance, fumarase is distributed between cytosol and mitochondria despite being processed by the mitochondrial matrix signal peptidase. Import of the unfolded protein seems to be incomplete, allowing retrograde movement of mature fumarase into the cytosol (Stein et al., 1994). Targeting of the *LEU4* gene product depends on the transcriptional starting point. Two of four major transcription starting sites are located downstream from this ATG causing cytosol localisation, while the longer transcripts, started upstream of the ATG, are translocated into the mitochondria (Beltzer et al., 1988). Lee et al. (Lee et al., 2000) found a cleavable N-terminal signal on the *CIT2* gene, which might act as ambidextrous organelle targeting signal to either peroxisome and mitochondria. Small and McAlister-Henn demonstrated the existence of a secondary targeting signal on the *MDH2* gene, which is also capable of directing the protein into the mitochondria in the absence of the primary targeting sequence (Small and McAlister-Henn, 1997). Furthermore, regulation of these targeting sequences is presently poorly understood, but it is very well conceivable, that the compartmentation of various enzymes is dependent on the cultivation conditions.

Although transport proteins are currently subject of ongoing research, their identification is far from being completed. Currently, about 500 genes have been identified to likely code for transporter proteins, many those are associated with cellular import and vesicular transport. Of these transporters more than half are still uncharacterized (Paulsen et al., 1998). Only 83 ORFs are associated with mitochondrial transport, of these, 20 have no function assigned and 27 are protein transport facilitators (http://www.mips.biochem.mpg.de/proj/yeast/catalogues/funcat/.. http://www.proteome.com). This leaves about 36 characterized mitochondrial carrier proteins for the metabolite translocation (el Moulalij et al., 1997). Even less information is available on the transcription regulation and the kinetics of these transporters (Palmieri et al., 2000).

Nevertheless, the knowledge of the compartmental distribution of the yeast’s metabolism is required for improved models, which will have the capability of predicting the result of genetic modifications or the cell’s behaviour under dynamic environmental conditions, since metabolites located in different compartments will provide precursors for distinct pathways. Furthermore, compartmentation will cause separate pools of the cofactors NAD(P)(H) and ATP/ADP for each organelle due to their inability to freely cross the
intracellular membrane. Several models have been put forward to at least partially address compartmentation in yeast. Nissen et al. (Nissen et al., 1997) were the first, who included a mitochondrial compartment in their metabolic model of anaerobic growth. However, compartmentation was limited to a few essential metabolites, but excluded e.g. ATP. The same holds for the work of Granström et al. (Granström et al., 2000) and Ostergaard et al. (Ostergaard et al., 2001). Compartmentation was restricted to NADPH in the work by Wahlbom et al. (Wahlbom et al., 2001).

The presented work attempts to give an overview of the current state in research into the compartmentation of the metabolism of S. cerevisiae. Based on this literature review, a coherent compartmented network model will be extracted for aerobic growth of yeast on either of three carbon substrates, namely glucose, ethanol, or acetate with ammonia as the nitrogen source. The discussion will focus primarily on the localisation of the enzymes involved in the catalytic steps of the central metabolism, this will be combined with the information currently available on transport proteins for metabolites across intracellular membranes. Of the many organelles in yeast, only the mitochondria and, to a certain degree, the peroxisome will be considered in the present discussion of the metabolic pathways in Saccharomyces cerevisiae.

Compartmental localisation of pathways

Glycolysis, pentose phosphate cycle, and NH₃-assimilation

Although S. cerevisiae is one of the most intensively studied organisms, many questions remain unsolved. One of these is the localisation of various enzymes. The best described pathway is probably the glycolysis, which is located exclusively in the cytosol, as is the complete pentose phosphate cycle. However, most pathways of the central metabolism are localised over several compartments, often with isoenzymes in two or even three compartments catalysing the same reaction in a different context. Good examples are the glutamate dehydrogenase (NADP) (GDH1, GDH3) and the GS-GOGAT system, the essential reactions involved in the metabolism during growth on ammonia as nitrogen source, which make the ammonium ion metabolically available through formation of either glutamate or glutamine. Of these enzymes, the Gdh1p is cytosolic (Perlman and Mahler, 1975) and Gdh3p is believed to be nucleic (Camardella et al., 1975). But as in many other cases, the localisation of the GOGAT system is unknown, and due to this lack of information a placement has been assumed for modelling purposes.

TCA cycle

The TCA cycle is the main source of the reducing equivalents NADH and FADH₂ in the mitochondria, and thus the main energy source during respiratory growth. It further provides anabolic precursors such as oxoglutarate and oxaloacetate. All enzymes required for this 9-step oxidation of acetyl-CoA can be found in the mitochondria. Isoenzymes of citrate synthase, isocitrate dehydrogenase, and malate dehydrogenase have been identified in the peroxisome and with the exception of citrate synthase also in the cytosol; a second aconitase was identified in the cytosol (Przybyla-Zawislak et al., 1999). The fumarate
hydratase protein (Fum1p) has been characterized in both cytosol and mitochondria due to its dual targeting sequence (Stein et al., 1994).

The first TCA cycle step catalyses the condensation of oxaloacetate and acetyl-CoA to form citrate. The citrate synthase isoenzymes are encoded by three genes, CIT1, CIT2, and CIT3, with Cit1p and Cit3p being mitochondrial (Jia et al., 1997; Suisa et al., 1984) and Cit2p peroxisomal (Lewin et al., 1999). The mitochondrial presence of Cit1p is essential for growth on acetate, but its catalytic activity is not required (Vélot et al., 1999). CIT1 deletion mutants can be rescued by over-expression of Cit3p or targeting of Cit2p to the mitochondria. The peroxisomal Cit2p is essential during growth on oleate, but is not required for acetate or ethanol metabolism. The TCA cycle is closed within the mitochondrion, through the aconitase Aco1p, isocitrate dehydrogenase Idh1p&2p, or alternatively Idp1p, the oxoglutarate dehydrogenase complex (Kdg1&2p + Lpd1p), the succinate CoA ligase (Lsc1&2p), the succinate dehydrogenase complex (Sdh1,2,3&4p), the fumarase Fum1p, and finally malate dehydrogenase Mdh1p.

Of the peroxisomal enzymes, isocitrate dehydrogenase Idp3p is required for the catabolism of polyunsaturated fatty acids, but not for growth on oleate (van Roermund et al., 1998). Of the three ORFs encoding for malate dehydrogenase, the peroxisomal Mdh3p is only present during β-oxidation of fatty acid, but is then essential (van Roermund et al., 1995), while the mitochondrial Mdh1p and the cytosolic Mdh2p are constitutively expressed. Summarising, the peroxisomal isoenzymes have, if at all, only a limited function during growth on glucose, ethanol, or acetate.

For provision of anabolic precursors and their subsequently required replenishing, the TCA cycle intermediates, most notably oxoglutarate and oxaloacetate, are required in the cytosol. For oxoglutarate, this requires the export out of the mitochondria either directly or as its TCA cycle predecessor citrate, aconitase, or isocitrate, which then have to be converted via the cytosolic TCA cycle isoenzymes. The second option requires an active cytosolic aconitase and/or isocitrate dehydrogenase. Similarly, oxaloacetate can be generated from fumarate through the fumarate hydratase and malate dehydrogenase in both the mitochondria or the cytosol. Furthermore, the cytosolic conversion of fumarate to malate and aspartate catalysed by the three cytosolic enzymes Fum1p, Mdh2p, and Aat2 allows the one-compartmental regeneration of amino-group donor aspartate, which is required in the cytosolic synthesis of arginine, AICAR, and AMP.

Pyruvate branch point

Pyruvate is a central metabolite at the interface between glycolysis, TCA cycle, and C2-metabolism. During respiratory growth on glucose, about 60% of the pyruvate is transformed via the mitochondrial pyruvate dehydrogenase complex or the pyruvate decarboxylase. A further 15% is required replenishing biosynthetic intermediates withdrawn form the TCA cycle and proceeds via the cytosolic pyruvate carboxylase isoenzymes encoded by PYC1 and PYC2 (Gombert et al., 2001). Flikweert et al. clearly demonstrated the need for the decarboxylase to provide the extra-mitochondrial acetyl-CoA (Flikweert et al., 1999); all its three isoenzymes Pdc1p, Pdc5p, and Pdc6p are found in the cytosol. At least 10% of the pyruvate has to be metabolised via this so-called pyruvate dehydrogenase bypass, which proceeds via the acetaldehyde dehydrogenase and
the acetyl-CoA ligase to yield acetyl-CoA in the cytosolic compartment. The remaining pyruvate is consumed for the synthesis of alanine, leucine, valine, and isoleucine.

During growth on ethanol and acetate, the general direction in the pyruvate branch point reverses; oxaloacetate (OAC) is decarboxylated via the PEP carboxykinase (Pck1p) to yield phosphoenol pyruvate, which then is dephosphorylated by the pyruvate kinase (PYK1 or PYK2) to form pyruvate. All these proteins are cytosolic (figure 1, pathway A). Yeast further harbours a mitochondrial decarboxylating malate dehydrogenase (malic enzyme, Mae1p), which also can provide pyruvate from malate (figure 1, pathway D). Boles et al. showed that despite their different localization either the MAEI or the PYK1 gene is sufficient for the growth on ethanol and acetate (Boles et al., 1998). Addition of small amounts of alanine or pyruvate rescued the double mutant, indicating that it was indeed the inability to form pyruvate for anabolic purposes, which caused the phenotype. It also demonstrated that pyruvate can be transported between the two compartments. The small energetic differences between the pathways A and D are hard to quantify with conventional metabolic flux analysis, since not much is known about the cost of transport involved in (A) or the energetic differences in mitochondrial NADH and NADPH in (D). Pathway (A) was chosen for modelling, assigning a small cost of one translocated proton for each transport step for the formation of mitochondrial pyruvate. Two further reaction can link oxaloacetate and pyruvate: The pyruvate carboxylase (Pyc1p or Pyc2p) catalyses the direct carboxylation of pyruvate to oxaloacetate, but is not able to facilitate the reverse direction (figure 1, pathway B). And secondly, the malic enzyme is capable of directly forming pyruvate from oxaloacetate (www.expasy.ch/cgi-bin/enzyme-search-ec), but is likely to require higher OAC concentrations than found in the mitochondria for this step to proceed (figure 1, pathway C).

**Figure 1:** possibilities of pyruvate formation from oxaloacetate in the mitochondria (black arrows) and the cytosol (grey arrows)

Control of the pyruvate branch point enzymes is closely related to the metabolism of ethanol, which is converted to acetaldehyde and acetate via alcohol dehydrogenase and
acetaldehyde dehydrogenase. For both, different isoenzymes exist. All characterized ethanol dehydrogenases in yeast are NAD dependent, Adh1p, Adh2p, Adh4p, and Adh5p are cytosolic enzymes, only Adh3p is mitochondrial. For the five isoenzymes of the acetaldehyde dehydrogenase, only the mitochondrial Ald5p and the cytosolic Ald6p are strictly NAD dependent. Ald2p, Ald3p, and Ald4p have a preference for NADP as cofactor. The cytosolic Ald2p and Ald3p are induced as stress response, the mitochondrial Ald4p is important, but not essential for growth on ethanol (Wang et al., 1998), thus these isoenzymes are unlikely to play a significant role under steady state chemostat cultivation.

Glyoxylate cycle

The glyoxylate pathway fulfills the anaplerotic requirements of the TCA cycle for growth on C₂-compounds and requires seven enzymes, citrate synthase, aconitase, isocitrate lyase, succinate dehydrogenase, fumarase, malate dehydrogenase, and malate synthase (figure 2). In eukaryotic cells the glyoxylate cycle enzymes are commonly located in the peroxisomes. Peroxisomal isoenzymes have been identified in Cit2p and Mdh3p (Lewin et al., 1999; McAlister-Henn et al., 1995), however the isocitrate lyase, encoded by ICL1 was only found in the cytosol in *S. cerevisiae* (Taylor et al., 1996). The highly homologous ICL2 gene encodes a glucose repressed mitochondrial 2-methyl-isocitrate lyase which does not act on isocitrate (Luttik et al., 2000). The produced glyoxylate is acetylated by the malate synthase encoded by MLS1 and DAL7. While Dal7p is cytosolic, Mls1p is reported to be peroxisomal, but its peroxisomal import is not required for growth on acetate (Brocard et al., 1996). The next step in the cycle is catalysed by the malate dehydrogenase encoded by *MDH1*, *MDH2*, and *MDH3*. Mdh3p is localized in the peroxisome and constitutes 10% of the total malate dehydrogenase activity under acetate growth. Its deletion mutant exhibits reduced growth on mineral medium with acetate as carbon source (Steffan and McAlister-Henn, 1992). However, the mainly cytosolic Mal2p is responsible for 65% of the total malate dehydrogenase activity in acetate grown cells, and the corresponding deletion mutant is unable to grow on neither ethanol nor acetate (Minard and McAlister-Henn, 1991).

![Glyoxylate cycle diagram](image)

**Figure 2:** compartmentation of the glyoxylate cycle reactions: occurrence of the enzymes in the mitochondria (black arrows) and the cytosol (grey arrows): citrate synthase (1), aconitase (2), isocitrate lyase (3), succinate dehydrogenase (4), fumarase (5), malate dehydrogenase (6), and malate synthase (7)

Summarising, the central steps of the glyoxylate cycle seem to be located in the cytosol rather than in the peroxisome. This corresponds with the peroxisome proliferation solely on oleate grown cells; growth on C₂-compounds seems not to require an additional activity in the peroxisomal compartment. The conversion of the formed succinate to fumarate involves a mitochondrial step and thus
trafficking of these metabolite across the mitochondrial membrane. All remaining steps can be catalysed in the cytosol, assuming the existence of a cytosolic citrate synthase and aconitase, a point which will be discussed later.

**Anabolic pathways**

While the vast majority of anabolic reactions proceed within the cytosol, there are a number of notable exceptions. Of all amino acids, only aspartate can be synthesised in all three compartments, the cytosol, the mitochondria, and peroxisome, allowing it a function as universal amino-group donor. The aspartate aminotransferase Aat1p is located in the mitochondria, while the AAT2 product is located in the peroxisome during growth on oleate but in the cytosol during growth on glucose (Verleur et al., 1997). Glycine production from serine can occur in either the cytosol (Shm2p) or the mitochondria (Shm1p) (McNeil et al., 1994), however this synthesis pathway is coupled to the C1-metabolism. The formation of glycine from threonine through the threonine aldolase (*GLY1*) is thought to be located exclusively in the cytosol. Two putative alanine amino transferases have been discovered in the yeast genome, of which the YDR111c is predicted to be cytosolic, while YLR089c is likely to be a mitochondrial enzyme.

The branched chain amino acid synthesis can proceed partly via isoenzymes in both compartments. The synthesis of 2-oxoisovalerate, valine, and isoleucine can take place completely in the mitochondria (Kispa et al., 1996), but parts of these pathways have been also observed in the cytosol. Notably, the branched chain amino acid transaminase has two isoenzyme, Bat1p & Bat2p, in the mitochondria and in the cytosol respectively, which both catalyse the amination of oxoisovalerate to valine and the final steps in the synthesis of leucine and isoleucine (Kispal et al., 1996). The first step of the leucine synthesis from oxoisovalerate is encoded by LEU4, the localization of its product depends on the transcriptional regulation and can be either mitochondrial or cytosolic (Beltzer et al., 1988). But since the intermediate steps have not been identified in the mitochondria, it is sensible to assume a cytosolic leucine synthesis for modelling purposes.

All other reactions of the amino acid synthesis pathways are exclusively located in one compartment, mostly in the cytosol. One mitochondrial pathway is the transformation from glutamate to ornithine, which proceeds via five catalytic steps (Arg2p, Arg5,6p, Arg8p, and Arg7p), all of which can be found in the mitochondria (Jauniaux et al., 1976). Of the eight steps of the α-aminoacidipate pathway to lysine, the first is located in the nucleus (Chen et al., 1997), the following three are located in the mitochondria, the next three are again cytosolic (Bhattacharjee, 1985), and the last lysine forming step is reported to be peroxisomal (Geraghty et al., 1999). Conflicting results are given for the third and eighth step, which are reported to be either peroxisomal (Geraghty et al., 1999) or cytosolic (Bhattacharjee, 1985). The nucleus and the cytosol can be treated as continuum, since the nucleic membrane is easily permeable for smaller molecules even without dedicated transport proteins.

Further pathways with cytosolic and mitochondrial localization are found in the C1-metabolism and the synthesis of phospholipids. The transformation of formyl-THF via methenyl-THF to methylene-THF can proceed in both compartments (Shannon and Rabinowitz, 1986). Phosphatidate (PHD) also can be elongated to CMP-diacylglycerol,
PHD-serine, and PHD-ethanolamine via isoenzymes located in either cytosol or mitochondria (Dzugasova et al., 1998; Kohlwein et al., 1988; Kuchler et al., 1986). Single step transformations with dual localisation are the ethanol dehydrogenase and the glycerol-3-phosphate dehydrogenase (http://mips.gsf.de/proj/yeast/CYGD/db) (figure 3). The reaction from glycerol-3-phosphate to dihydroxyacetone phosphate can also be catalysed by Gutp, which is not located in the mitochondrial matrix, but in the inner membrane. Instead on acting on NAD, it is coupled directly to the ubiquinone pool of the respiratory chain and the catalytic site of this membrane bound enzyme Gut2p is cytosolic (Grauslund and Rønnnow, 2000).

**Peroxisomes**

The peroxisomes play a crucial role during the catabolism of fatty acids. They are also generally believed to harbour the enzymes of the glyoxylate cycle in eukaryotic cells. However, the evidence found for *S. cerevisiae* points to a minor role of the peroxisome during growth on substrates other than fatty acids. It was in fact only the discovery of peroxisomal proliferation on oleate by Veenhuis et al., that established the existence of this organelle in this yeast beyond doubt (Veenhuis et al., 1987). Therefore, practically all peroxisomal enzymes attributed to this compartment were found in yeast cells grown on oleate. In several cases the growth conditions were not reported in sufficient detail to allow the determination of the carbon source, e.g. Geraghty et al. (Geraghty et al., 1999). Therefore, the peroxisome was not included in the modelling of the metabolic networks of yeast grown on glucose, ethanol, or acetate with ammonia as nitrogen source.

**Cofactor regeneration in different compartments**

The division of the cell’s intracellular space into several compartments raises the need for distinct pools of acetyl-CoA/CoA, NAD+/NADH, NADP+/NADPH, and ADP/ATP which can not pass the intracellular membranes, and for which no direct transporters are present (van Roermund et al., 1995). Therefore, regeneration of the cofactors has to proceed either via distinct reactions in each compartment or through the operation of a shuttle system, whereby a net-transport is accomplished through the operation of a cofactor dependent, reversible reaction in different compartments combined with the transport of the involved metabolites.

NADPH can be mainly provided in the cytosol through the glucose-6-phosphate dehydrogenase (Zwf1) and phospho-gluconate dehydrogenases (Gnd1p & Gnd2p) as part of the pentose-phosphate (PP) cycle. Inclusion of both the glycolysis and the PP pathway gives a metabolic model the degree of freedom required to provide exactly the amount of cytosolic NADPH needed. While the PP-pathway is the major cytosolic NADPH source for growth on glucose, for ethanol it will be the cytosolic acetaldehyde dehydrogenases, all of which are, although not strictly, NADP specific. However, the low flux through this enzyme would only provide a small share of the total NADPH demand during growth on
glucose and acetate. During growth on acetate, a cytosolic isocitrate dehydrogenase must provide the required NADPH.

Mitochondrial NADPH is generated by the NADP dependent isocitrate dehydrogenase (Idp1p), activity of both, the NAD dependent isocitrate dehydrogenase consisting of the subunits Idh1p and Idh2p, and Idp1p allows for the exact provision of the anabolically required NADPH for growth on glucose, ethanol and acetate. A further source during growth on ethanol is the NADPH dependent acetaldehyde dehydrogenase (ALD4), which has a dual cofactor specificity and is required for normal growth on ethanol, and is only expressed after glucose depletion (Boucherie, 1985). Another potential source of mitochondrial NADPH is the malic enzyme, but the lack of a phenotype for a mael∆ mutant indicates its contribution as non-essential (Boles et al., 1998).

Shuttle systems

The lack of specific transporters of e.g. NAD(H) or glutamate requires the use of shuttle systems. They provide an indirect mechanism of transport by donating the moiety in question to another molecule, which is then transported across the membrane and subsequently undergoes the reverse steps in the other compartment. They are thus capable of transferring the redox potential of NADH between the cytosol and the mitochondria, despite the lack of NAD(H) transporters. Bakker et al. demonstrated that the isoenzymes of alcohol dehydrogenase are involved in the indirect NADH transport across the mitochondrial membrane (figure 3A) (Bakker et al., 2000). Other shuttle systems for which both isoenzymes are confirmed the malate/oxaloacetate shuttle via malate dehydrogenase (figure 3C), and the malate/pyruvate shuttle via the malic enzyme, pyruvate carboxylase, and malate dehydrogenase (figure 3B) (Bakker et al., 2001). In all cases the redox potential of NADH in the cytosol is transported into the mitochondrion via a reduced metabolite, which is subsequently oxidized providing thus mitochondrial NAD(P)H and transported back in its oxidised form to the cytosol. The GOH/DHAP shuttle involving the membrane bound glycerol-3-phosphate dehydrogenase Gut2p does provide a possibility to deliver the redox potential of NADH to the quinone pool.

Strictly spoken not a cofactor, but from a modelling perspective acting similarly, aminogroups are required in many of the anabolic pathways and need also to be transported across the intracellular membranes. The three donor systems aspartate/fumarate, glutamine/glutamate and glutamate/oxoglutarate all operate in the cytosol, but only the latter has been identified as operating in the yeast mitochondria.

Regeneration of glutamate during growth on excess ammonia occurs via the glutamate dehydrogenase, whose relevant isoenzymes Gdh1p and Gdh3 are cytosolic. Assuming a cytosolic GOGAT system, as suggested by the lack of a phenotype of the GDH knock-out mutant on ammonia (Miller and Magasanik, 1990), a functional metabolic network requires a shuttle system for the provision of mitochondrial glutamate. The required isoenzymes are available for an aspartate/oxaloacetate shuttle or for a shuttle system involving the three transaminations catalysed by the branched chain amino acid transferase (BAT1 & BAT2), using e.g. the metabolite couple valine/oxoisovalerate. These options would require the mitochondrial export of, respectively, oxaloacetate and oxoisovalerate and the import of aspartate and valine (figure 3). Kispal et al. observed not only an
auxotrophy for the branched chain amino acids (Leu, Ile, Val) in a bat1Δbat2Δ deletion mutant, but also severe growth reductions on glucose when these amino acids were supplied, an indication, that these enzymes have an additional metabolic role. Excluding the possibility of imperfect targeting, these results also indicate, that trafficking of the precursor seems to be possible, since neither the deletion of the cytosolic Bat2p nor the mitochondrial Bat1p were strictly auxotroph (Kispal et al., 1996).

![Diagram of metabolic pathways involving NADH and amino-acids](image)

**Figure 3:** shuttle systems for NADH and amino-groups; NADH shuttles: ethanol/acetaldehyde shuttle via the alcohol dehydrogenases Adh1p, Adh2 & Adh3p (A) - malate/pyruvate shuttle via pyruvate carboxylase Pyc1/2p, malate dehydrogenase Mdh2p and the malic enzyme Mae1p (B) - malate/oxaloacetate shuttle via malate dehydrogenases Mdh1&2p (C) - malate/aspartate shuttle via malate dehydrogenases Mdh1&2p and aspartate transferases Aat1&2p (D); provision of redox potential to the quinone pool via the glycerol-3-phosphate / dihydroxyacetone phosphate shuttle using the glycerol-3-phosphate dehydrogenases Gdp1 and Gut2p (E); the amino-group shuttle using the branched chain amino acids transferases Bat1p and Bat2p with transport of oxoisovalerate and valine (F)

** Provision of acetyl-CoA

The acetyl-CoA is the basic donor of a C₃-carbon-skeleton and is used throughout the metabolism, providing the entry point into the TCA-cycle via the citrate synthase and the glyoxylate cycle via the malate synthase, but also e.g. the basic building blocks of the fatty acid synthesis. A network based on two intracellular compartments requires either two distinct sources of acetyl-CoA, since the mitochondrial membrane is impermeable to acetyl-CoA, or alternatively, a possibility to transport it into the mitochondria.

For growth on glucose, the only significant sources of acetyl-CoA are the pyruvate dehydrogenase complex and acetyl-CoA synthase. The multi-protein complex of pyruvate dehydrogenase is mitochondrial and provides acetyl-CoA for the mitochondrial citrate synthase. However, the results of Flikweert et al. (Flikweert et al., 1999) demonstrated the necessity of pyruvate decarboxylase for growth on glucose. The amount of the required ethanol or, alternatively, acetate points towards the provision of the cytosolic acetyl-CoA via this pathway (figure 4). The transformation of acetaldehyde to acetyl-CoA can proceed
via one of the five identified acetaldehyde dehydrogenases and one of the two acetyl-CoA synthases (Acs1p and Acs2p). Based on the findings of Flikweert et al., it is logical to assume that these steps occur in the cytosol for at least one combination of both enzymes. Of the acetaldehyde dehydrogenases Ald2p, Ald3p, and Ald6p have been identified as being cytosolic (Wang et al., 1998). While Ald2p and Ald3p are stress response proteins, the reported requirement of Ald6p is disputed (Meaden et al., 2000; Wang et al., 1998). Since the acetyl-CoA synthase gene ACS1 is repressed by glucose (van den Berg et al., 1996), the second step is expected to proceed via ACS2 in glucose grown batch cultures. This would require the Acs2p to be cytosolic, which is supported by the absence of a targeting sequence (van den Berg and Steensma, 1995) and the findings of de Jong-Gubbels (de Jong-Gubbels, 1998).

\[
\text{Glycolysis} \quad \downarrow \quad \text{ETOH}_{\text{cyt}} \\
\quad \downarrow \quad \downarrow \\
\text{PYK1&2} \quad \downarrow \quad \text{ADH1&2} \\
\downarrow \quad \downarrow \\
\text{PYR}_{\text{cyt}} \quad \rightarrow \quad \text{ACTAL}_{\text{cyt}} \quad \rightarrow \quad \text{ACTAL}_{\text{mit}} \\
\quad \downarrow \quad \downarrow \quad \downarrow \\
\text{PDC1.5&6} \quad \text{ALD6} \quad \text{ALD4} \\
\downarrow \quad \downarrow \quad \downarrow \\
\text{PYR}_{\text{mit}} \quad \text{ACT}_{\text{cyt}} \quad \rightarrow \quad \text{ACT}_{\text{mit}} \\
\quad \downarrow \quad \downarrow \\
\text{PDA1+PDB1} \quad \text{ACS1&2} \quad \text{ACS1?} \\
\quad + \text{LPD1+LAT1} \\
\quad + \text{PDX1} \\
\downarrow \\
\text{ACCoA}_{\text{mit}} \quad \text{ACCoA}_{\text{cyt}} \quad \text{ACCoA}_{\text{mit}}
\]

**Figure 4:** provision of acetyl-CoA for mitochondria and cytosol under different substrates (underlined) through cytosolic (grey) and mitochondrial (black) reactions, transporters are shaded.

For growth on ethanol or acetate, the acetyl-CoA synthase is the only major source of acetyl-CoA, and with this the only provider of C_2-carbon bodies to close the upper part of the mitochondrial TCA cycle. It is, therefore, sensible to assume production of mitochondrial and cytosolic acetyl-CoA with two distinctly located acetyl-CoA synthases enzymes. However, three other options involving the cytosolic formation of acetyl-CoA will also be considered in the following paragraphs: 1.) the possibility of a cytosolic citrate synthase, which would close the TCA cycle not in the mitochondria, but in the cytosol, 2.) the existence of a shuttle system involving oxoisovalerate, and 3.) the carnitine shuttle, which is frequently found in eukaryotic cells.

Two isoenzymes of the acetyl-CoA synthase are found in yeast. While all evidence points towards a cytosolic localization of the constitutively expressed Acs2p, the evidence is less clear for the Acs1p. De Virgilio (De Virgilio et al., 1992) pointed out, that the ACS1 gene possesses two distinct ATG starting codons with different targeting sequences for
mitochondria and microsomes. Klein and Jahnke (Klein and Jahnke, 1968) indeed reported the dependance of the acetyl-CoA synthase localisation on the growth conditions in aerobic cultures. During batch growth on glucose, they found two distinctly localized fractions depending on the age of the culture, the earlier one associated with the microsomal and the later with the mitochondrial fraction. However, these results could be explained by the reported association of Acslp with the outside of the peroxisomal membrane during growth on oleate (de Jong-Gubbels, 1998). While a mitochondrial presence of acetyl-CoA synthase during growth on ethanol or acetate can not be completely discarded, the findings of de Jong-Gubbels et al. weigh against this hypothesis. Furthermore, a mutant expressing only ACS1 (acs2Δ) can grow on glucose under aerobic, glucose limited conditions, where no glucose repression occurs (van den Berg et al., 1996). This indicates a cytosolic localization of the enzyme, which might, however, not be exclusive.

The second conceivable option would be a cytosolic citrate synthase, which provides an alternative to the import of acetyl-CoA into the mitochondria. The catalysed condensation of oxaloacetate and acetyl-CoA would provide cytosolic citrate, which can be transformed in the cytosol via aconitate into isocitrate, or even further into oxoglutarate. Cytosolic isoenzymes have been identified for aconitate (ACOI) and isocitrate dehydrogenase (IDP2). Either citrate, isocitrate, or oxoglutarate would have to be imported into the mitochondria via a tricarboxylate carrier protein, which has been identified in the form of the citrate transport protein (CTP). However, all of the identified three isoenzymes of the citrate synthase are claimed to be either mitochondrial or peroxisomal (Jia et al., 1997; Lewin et al., 1999) with the appropriate targeting sequences identified (Lee et al., 2000). During their localization studies, Lewin et al. demonstrated that the majority of citrate synthase activity cosedimented with the mitochondrial cytochrome-c oxidase or the peroxisomal catalase marker enzymes. However, about 20% was recovered in the cytosolic fraction, far more than could be expected from rupture of mitochondria or peroxisomes (Lewin et al., 1999). Vélot et al. report the same relative amount of activity in the wild type, an even higher percentage is found for a cit1Δ deletion mutant (Vélot et al., 1999). During their studies of peroxisomal enzymes McCammon et al. could only recover 70% of the citrate synthase activity in the crude peroxisomal-mitochondrial pellet (McCammon et al., 1990). Furthermore, expression of the CIT3 gene was found in glucose limited chemostat cultures despite the absence of peroxisomes under these conditions (J. Pronk, personal communication). These results indicate, that a significant amount of citrate synthase activity can be found in the cytosol although the protein is mainly located in the organelles.

A third alternative would be the use of a shuttle system involving isopropyl malate synthase. The two isoenzymes encoded by LEU4 and LEU9 catalyse the condensation of oxoisovalerate and acetyl-CoA to 2-isopropyl-malate and are located in the cytosol and the mitochondria, respectively. This possibility is supported by the growth defect of the leu4Δ mutant on non-fermentable substrates, but almost normal growth on glucose (Casalone et al., 2000). However, no carrier has yet been identified neither for 2-isopropyl-malate nor for oxoisovalerate, but localisation of the branched chain amino acid synthesis pathways strongly is suggesting the existence of at least the latter one.
The fourth route for the supply of mitochondrial acetyl-CoA is the carnitine shuttle system, requiring two distinctly located acetyltransferases and the appropriate transport protein. In yeast, the \textit{CRC1} gene has been identified as encoding an inner mitochondrial membrane carrier protein, which acts as antiport on propyl- and acetyl-carnitine (Palmieri et al., 1999). The necessary carnitine O-acetyltransferases have been found on the outer side of the mitochondrial membrane (Yat1p) and in the mitochondria and the peroxisome (Cat2p). Both enzymes are expressed during acetate, but not during ethanol consumption during a glucose pulse (van den Berg et al., 1998). This might indicate two distinct catabolic routes for acetate and ethanol, but leaves open the question of the mitochondrial import of the two-carbon-skeleton during ethanol growth. Furthermore, \textit{S. cerevisiae} seems to be unable to synthesize carnitine \textit{de novo} and requires a plasma membrane carnitine transporter, Agp2p for its uptake from the medium (van Roermund et al., 2000). Hence, a functional carnitine shuttle seems unlikely in yeast grown on mineral medium on glucose, ethanol, or acetate.

This leaves three feasible options for the mitochondrial import of the C\textsubscript{3}-carbon body. Option one requires the transport of either acetaldehyde or acetate into the mitochondria for, respectively, ethanol and acetate as substrate. A substantial amount of acetyl-CoA synthesis would occur in the mitochondria, while taking place exclusively in the cytosolic for option two and three, which requires the additional export of mitochondrial produced ATP. Import of a TCA cycle intermediate, as required by the second option, needs furthermore transport systems for the export of the C\textsubscript{4}-compound, e.g. oxaloacetate, and for the import of the tricarboxylate. Both could be combined in one antiporter, exchanging, e.g., malate against citrate. All three options are energetically slightly different.

**Intracellular transport of metabolites**

The compartmentation of the cell's interior requires the classification of the metabolites according to their ability to cross the intracellular membranes. Only a few small molecules can diffuse freely across biological membranes, all other require facilitating proteins. The latter can either diffuse through pore protein or can be actively transported by dedicated transport proteins. The absence of a suitable carrier protein confines the metabolite to its original compartment.

The number of carrier proteins in \textit{S. cerevisiae} is estimated in the order of several hundred. About 340 ORFs have been assigned as encoding for recognized and putative transporters (Paulsen et al., 1998). Of these, only about 35 belong to the family of mitochondrial membrane transporters, most other facilitate the transport across the cell membrane. Thus, only little is know about possible metabolite transport between compartments.

A relevant metabolic network requires further the information of the energetics of each included step, thus also the applied transport mechanism for translocation steps. Translocation via free or facilitated diffusion does not consume energy in the form of ATP or the reduction of the H\textsuperscript{+} gradient, except in the case in which the solute is charged and transport would effect the electric component of the proton motive force $\Delta \psi$. Insofar, it is
energetically comparable to an antiport facilitating the electroneutral exchange of two components. Generally, the ubiquitous molecules oxygen, carbon dioxide, and water can be transferred between the compartments without consumption of energy, but most other components are likely to require some form of energy input for their translocation. This can occur e.g. via a proton symport into or a proton antiport out of the mitochondria which uses the proton-motive force to drive the reaction. These mechanisms were used in the many cases, where specific data was not available.

**Generally required intracellular transporters**

Depending on the main carbon substrates, different sets of transporter proteins are required. Nevertheless, a majority of transporter proteins is always needed. Ever-present under aerobic growth conditions are the transport proteins involved in the ATP regeneration. The F₀F₁-ATPase uses the proton-motive force generated by the respiratory chain to catalyse the mitochondrial phosphorylation of ADP to ATP at the inner side of the inner mitochondrial matrix. The F₀F₁-ATPase complex is made up of sixteen distinct subunits (Arnold et al., 1999), thus accounting for almost half of the identified mitochondrial carrier proteins. The ATP can subsequently be exported from the mitochondria via an ATP/ADP carrier protein (Aac1p, Aac2p, and Aac3p), which acts as antiport for ATP and ADP. The required phosphate group can enter the mitochondria via the proton symport mechanism of the mitochondrial phosphate carrier protein (Mir1p) (Nielsen, 1998). The transport is electrically neutral, the phosphate ion H₂PO₄⁻ is translocated together with one proton H⁺ (Guérin, 1991).

**Mitochondrial amino acid synthesis requires the transport of oxoisovalerate, isoleucine, ornithine, oxoadipate (a lysine precursor) and possibly valine into the cytosol.** For these, only an ornithine carrier Ort1p and two oxodicarboxylate carriers Odc1p and Odc2p have been identified. The ornithine carrier is a proton antiporter (Palmieri et al., 1997); Odc1p and Odc2p function via a strict counter exchange mechanism, oxoglutarate, oxoadipate, but also adipate, glutarate, malate and citrate are suitable counter metabolites (Palmieri et al., 2001). Thus, a net-export would require a further transporter for one of these metabolites. In the absence of any identified carriers for isoleucine, lysine, valine, and oxoisovalerate, a proton antiport system is assumed for modelling purposes.

The mitochondrial synthesis of these amino acids requires the import of there precursors pyruvate, threonine, and glutamate into this compartment. Cytosolic cysteine formation is releasing oxobutyrate, which also has to be further metabolised in the mitochondria. None of the predicted carriers have so far been identified, a proton symport mechanism was, therefore, assumed for pyruvate, threonine, and oxobutyrate. Mitochondrial glutamate can be obtained through a transaminase reaction, e.g. the aspartate transaminase, which would shift the demand to a mitochondrial oxoglutarate, for which the Odc1p and Odc2p transporters have been identified. A coupling with the ornithine transport processes would also eliminate the need for an energy driven net-export. The mitochondrial localization of the threonine dehydratase further raises the need for an ammonia export from the mitochondria. Again no literature was found on the possible existence of such a carrier, and a proton antiport was used for the model.
As discussed earlier, transamination reactions in the mitochondria involving glutamate and oxoglutarate require the transport of an amino acid and its corresponding oxo-acid across the mitochondrial membrane. A shuttle system would export the oxo-acid and import the amino acid. While no transporters for amino acids are currently described, their existence can be readily assumed, due to the fact that proteins are assembled in the cytosol while several amino acids are exclusively produced in the mitochondria. Focussing on the oxo-acids instead, it becomes apparent that an exporter for mitochondrially produced oxoisovalerate must also exist to enable the cytosolic formation of leucine, while for oxaloacetate only a proton symport, thus a dedicated importer, is described in literature (Palmieri et al., 1999a).

The oxaloacetate carrier protein is encoded by OAC1 and the transport occurs as free acid driven by the pH-gradient across the mitochondrial membrane (Palmieri et al., 1999a). This indicates an import function of the oxaloacetate carrier, which is likely to fulfill the anaplerotic requirements of the TCA cycle. However, most of the TCA cycle needs for the formation of biomass are generally covered by antiports, as will be shown in the following paragraphs. Therefore, the role of the oxaloacetate carrier is limited to replenishing the TCA cycle intermediates lost due to cell dilution. This also helps to explains the finding, that the Oac1p is not essential and only the additional elimination of the dicarboxylate carrier gene DIC1 created a phenotype on non-fermentative carbon sources (Palmieri et al., 1999b).

Aspartate is further involved in the cytosolic transamination steps in the arginine, AICAR, and AMP synthesis, leading to cytosolic fumarate. Its regeneration occurs via fumarate hydratase, malate dehydrogenase, and aspartate aminotransferase and is likely to be located in the cytosol to minimize cross-membrane transfer. These reactions have thus been included in the model for both compartments, despite not being required for precursor provision.

Finally, the mitochondrial synthesis of ornithine requires glutamate, which can be provided via transamination of mitochondrial oxoglutarate. This step results in a depletion of the TCA cycle intermediates, and thus requires the operation of a mitochondrial importer, such as the OAC carrier.

**Intracellular transporters for growth on glucose**

For growth on glucose, the main entry point for C₇-bodies into the mitochondria is pyruvate, requiring a proton symport based import mechanism for either the acid or the ionized form. Both options should be tested in a metabolic network model. The anaplerotic requirements of the TCA cycle for C₇-skeletons are fulfilled by the cytosolic pyruvate carboxylase, providing oxaloacetate, a part of which requires import into the mitochondria. Fumarate hydratase (Fum1p), malate dehydrogenase (Mdh2p), and aspartate aminotransferase (Aat2p) have been identified in the cytosol in addition to their mitochondrial presence as discussed earlier. This enables any of the four metabolites (fumarate, malate, oxaloacetate, or aspartate) to be the point of translocation of the needed precursor into the mitochondria.
Chapter 4

The first TCA-cycle intermediate again required in the cytosol for the anabolism is oxoglutarate. Since the cell has a cytosolic aconitase Acolp as well as a cytosolic isocitrate dehydrogenase (Idp2c) export of either citrate, aconitate, isocitrate or oxoglutarate would fulfill the metabolic requirements. The choice of the transported metabolite does not cause a difference in the cell’s energetics, if the translocation follows the same transport mechanism for each of the compounds.

The existence of a functional citrate transporter in the mitochondrial membrane has been deduced from the ability of isolated yeast mitochondria to readily metabolise citrate. Oxygen consumption was observed for mitochondria obtained from cells grown on glycerol/lactate as well as on acetate. Rates compared favourably with the ones of succinate metabolism (Zhao and McAlister-Henn, 1996). The Ctp1p protein was later identified as a citrate carrier by Kaplan et al. (Kaplan et al., 1995). However, this protein seems to be an antiport system which facilitates the citrate uptake along with a proton in exchange for another tricarboxylate/H+ pair, isocitrate and phosphoenol pyruvate were identified as competitive inhibitors and are thus likely to also be exchanged. The choice for a citrate transporter would require a cytosolic aconitase and isocitrate dehydrogenase. Other antiports involved in the trafficking of oxoglutarate are the 2-oxodicarboxylate transporters Ocd1p and Ocd2p, which facilitate the exchange of oxoglutarate against oxoadipate or malate. No direct export protein for either the tricarboxylic acids or glutamate has yet been found. Since oxoglutarate export requires an anaplerotic reaction to refill the TCA cycle, the choice of an antiport system minimizes the cost of intracellular trafficking, the oxoglutarate/malate exchange was subsequently used for the model. In combination with an oxaloacetate transporter (OAC1), the system has sufficient freedom to alter the concentrations of the TCA cycle metabolites.

The small requirements of succinyl precursors can be neglected for modelling purposes. Thus, the next TCA cycle metabolite needed for anabolism is oxaloacetate and subsequently aspartate. The requirements for the C4-carbon skeletons can be directly fulfilled via the cytosolic aspartate aminotransferase (Aat2p) from cytosolic OAC, eliminating the need for membrane transport. This is supported by the fact, that the enzyme is located in the cytosol during growth on glucose, but is found in the peroxisome during growth on oleate (Verleur et al., 1997).

Transmitters for growth on ethanol and acetate

Generally, all transporter proteins required for glucose metabolism are also needed for C2-carbon sources. The additional glyoxylate cycle pathway and the different source of acetyl-CoA demand the existence of a number of additional carriers. The pyruvate transport could be considered an exception to these rules, if its provision as mitochondrial precursor for oxoisovalerate synthesis would be carried out via the malic enzyme. However, this can be equally well done via the cytosolic pyruvate kinase as shown by Boles et al. (Boles et al., 1998) and which was chosen for a metabolic model. Both pathways are approximately equivalent, requiring one proton for the import of either oxaloacetate or pyruvate into the mitochondria, the use of the malic enzyme would further transfer the redox potential of one NADH to NADP.
Different possibilities have been proposed for the import of the C₃-carbon-skeleton into the mitochondria. The first is based on a mitochondrial acetaldehyde dehydrogenase and a mitochondrial acetyl-CoA synthase, which would eliminate the need for any active transport for growth on ethanol, since acetaldehyde is highly volatile and likely to cross the membrane without a transporter protein (figure 5A). Growth on acetate would still require a mitochondrial acetate importer. The second possibility uses a cytosolic citrate synthase in the absence of a mitochondrial acetyl-CoA synthase. This would require the import of citrate, isocitrate, aconitate, or oxoglutarate into and the export of oxaloacetate or its precursors from the mitochondria (figure 5B). The mitochondrial tricarboxylate transport protein Ctp1p has been shown to facilitate the exchange of citrate and malate (Kaplan et al., 1995; Palmieri et al., 1972; Xu et al., 2000). The excreted malate can be reimported into the mitochondria via either the OAC carrier or the dicarboxylate carrier. The used of a acetyl-CoA shuttle in the third option would require transporters of oxoisovalerate and isopropyl-malate, but would with the exception of the first step be similar to figure 5A.

**Figure 5:** provision of C₃-bodies during growth on ethanol and acetate via the glyoxylate cycle with either a mitochondria (A) or cytosolic (B) citrate synthase. Compartmentation is indicated via grey (cytosolic) and black (mitochondrial) arrows, ovals indicate transporters out (arrow upward) or into (arrow downward) the mitochondria; antiports are indicated via a line between both exchanged metabolites.

The operation of the isocitrate lyase of the glyoxylate cycle requires cytosolic isocitrate. While the second option could directly provide it via a cytosolic aconitase, the first pathway demands the specific export of the isocitrate from the mitochondria into the cytosol, where it is split into glyoxylate and succinate (figure 5). This can also be accomplished by the tricarboxylate transport protein Ctp1p, now operating in the reverse direction compared with the previous case and exchanging isocitrate for malate. The produced succinate can only be further oxidized in the mitochondria, has thus to be funneled back into the mitochondria. This can be done through the fumarate-succinate antiporter encoded by *SFC1* (Pallota et al., 1999) which is essential for growth on ethanol.
and acetate. Again, the cytosolic malate produced by the glyoxylate cycle has to be reimported into the mitochondria.

The operation of an antiporter makes a special ex- and importer dispensable, while simultaneously minimizing the energetic costs of transportation. Active transport on ethanol and acetate is therefore only required of the net-import of the C2-body used in the energy generation in the TCA cycle, but hardly for anabolic purposes. Palmieri et al. demonstrated the requirement of the \textit{DIC1} gene for growth on ethanol and acetate (Palmieri et al., 1999b), it is therefore sensible to assume that the import actually proceeds via this enzyme during growth on a C2-carbon compounds, in contrast to the situation on glucose, where the OAC carrier seems to be predominant. Both options are energetically equivalent and can not be distinguished through metabolic flux analysis.

Oxaloacetate is consumed for gluconeogenesis via the cytosolic phosphoenol-pyruvate carboxylase during growth on ethanol and acetate and thus has to be either supplied directly through the cytosolic glyoxylate cycle or exported from the mitochondria. In the latter case, the function of the dicarboxylate carrier protein Diclp would shift from import to export. The operational direction of this phosphate/dicarboxylate antiport depends on the phosphate concentrations in both compartments. Hence, the active transport of phosphate into the mitochondria via Mirlp makes Diclp likely to also import the dicarboxylate. Its function is, therefore, physiologically identical with the OAC carrier.

**Differences for growth on ethanol and acetate**

Up to this point, ethanol and acetate grown cultures were considered together, with only minor differences between the metabolism of both substrates. However, significant differences were reported for various mutants grown on ethanol or acetate, which displayed distinct phenotypes for both substrates. Some of these observations will be discussed in the following paragraphs.

Two distinct responses were observed to the deletion of the \textit{ACSl} gene: the mutant was able to grow on ethanol, but not on acetate. If both, Acs1p and Acs2p, are located in the cytosol as proposed by de Jong-Gubbels (de Jong-Gubbels, 1998), this might be explained with the regulatory properties of the \textit{ACs2} gene, which is repressed by acetate and induced by ethanol (van den Berg et al., 1998). However, if both isoenzymes are located in distinct compartments, the mutation should be lethal with both substrates, unless the targeting of Acs2p is not exclusively mitochondrial and provides a residual activity in the cytosol.

Furthermore, large differences between the two carbon sources were observed on TCA cycle mutants. Przybyla-Zawislaw et al. constructed single deletion mutants of all TCA cycle enzymes and found ethanol to be the most permissive of all tested non-fermentable carbon sources, the least permissive being acetate (Przybyla-Zawislaw et al., 1999). Deletion mutants of isocitrate- \textit{(IDHI, IDH2)}, oxoglutarate- \textit{(KGD1, KGD2, LPD1)}, malate dehydrogenase \textit{(MDHI)}, and citrate synthase \textit{(CITI)} grew well on ethanol, but failed to grow on acetate. While all other mutants with the exception of \textit{aco1Δ} grew at
least slowly on ethanol, only the succinyl-CoA ligase deletion was allowing acetate growth (Przybyla-Zawislak et al., 1999). Kispal et al. had shown earlier that the presence but not the catalytic function of the Cit1p was required for growth of acetate (Kispal et al., 1989).

The mitochondrial NAD dependent isocitrate dehydrogenase consisting of the two subunits encoded by IDH1 and IDH2 is required for growth on acetate, while the double deletion mutant is still able to grow on ethanol. Zhao and McAlister-Henn demonstrated that even a cytosolic IDH is able to restore growth on acetate (Zhao and McAlister-Henn, 1996). This might also be caused by the stabilizing properties of IDH on mitochondrial mRNA. A similar observation in respect to the observed phenotypes was made for malate dehydrogenase. The mdh1Δ mutant’s growth rate was restored to near-parental level by a cytosol-directed mdh1 mutant protein on acetate and ethanol (Small and McAlister-Henn, 1997).

We are currently unable to sufficiently explain the reasons for these distinct phenotypes within the framework of a metabolic network. One possible explanation might be found in the different energy content of ethanol and acetate. The latter substrate provides significantly lower amounts of redox equivalents per mol substrate and requires additional ATP for transport. Other reasons are possibly a very different regulatory response on acetate, as can be found for the ACS2 gene, or the potential toxicity of acetate. The requirement of a structural but not functional equivalent of Cit1p for acetate demonstrates furthermore the importance of protein-protein interactions, which are outside the scope of a metabolic network.

Respiratory chain and ATP generation

During aerobic growth, the majority of ATP is regenerated through the operation of the respiratory chain, which facilitates the transfer of electrons from the adenine moieties to the final acceptor, oxygen via a system of four mitochondrial transmembrane complexes. During this process, protons are translocated from the mitochondrial matrix towards the space between the inner and outer mitochondrial membranes. The resulting proton-motive force is subsequently used to generate ATP via the F⁰F₁-ATPase. The respiratory chain located in the inner mitochondrial membrane can be fed by three distinct substrates: mitochondrial NADH, cytosolic NADH, and mitochondrial succinate which will donate its electron pair via FADH₂ (figure 6).

The yeast S. cerevisiae does not posses the common complex I of most eukaryotic cells capable of active proton transport. Instead, its two NADH dehydrogenases directly transfer two electrons from the NADH together with two protons to ubiquinone to form ubiquinol and NAD⁺ without further translocation of protons across the membrane (Bakker et al., 2001). This reaction can occur on either side of the inner mitochondrial membrane: cytosolic NADH is oxidized via Nd1p, the NADH dehydrogenases facing the outer side are encoded for by NDE1 and NDE2. Mitochondrial oxidation of succinate to fumarate proceeds via the succinate dehydrogenase complex (complex II), encoded by SDH1,
SDH2, SDH3, and SDH4. The FAD moiety is covalently bound to the complex and facilitates the reduction of ubiquinone as final step catalysed in the complex.

In all three cases the ubiquinol transfers the electrons to the complex III (cytochrome c-reductase), where two protons are excreted at the cytosolic side of the membrane. During the cycling of the second electron via the Q-cycle, two protons get translocated from the mitochondrial matrix to the inter-membrane space. The electrons are finally shuttled via the cytochrome c protein to the complex IV (Matsuno-Yagi and Hatefi, 1996). The complex IV (cytochrome c-oxidase) requires four reduced cytochrome c molecules to facilitate the reaction of O_2 and 4 H⁺ to two water molecules. During this process four further protons are translocated across the membrane per mol O_2 (figure 6).

\[ \text{Figure 6: the respiratory chain in Saccharomyces cerevisiae, required transporters are indicated by circles around the metabolite} \]

Thus a total of six protons is removed from the mitochondrial matrix and added to the inter-membrane space for each FADH₂ or NADH, mitochondrial or cytosolic, thus creating
the potential for six protons to be channelled back into the inner side. The number of 
translocated protons does not depend on the source of the initial electron source, in 
contrast to most other eukaryotic cells, e.g. *C. utilis*, which generate a potential of eight 
protons for each mitochondrial NADH and only six for cytosolic NADH.

The proton-motive force is used by the F₀F₁-ATPase, in *S. cerevisiae* a complex of 16 
different subunits located in the inner mitochondrial membrane. Here ATP is regenerated 
from ADP and phosphate. Ten translocated protons are thought to be channelled back into 
the mitochondria to regenerate three ATPs via the F₀F₁-ATPase (Stock et al., 1999), a ratio 
recently also found in *E. coli* (Jiang et al., 2001). The non-integral number of protons per 
ATP is still subject of discussions, but increasingly accepted (Arechaga and Jones, 2001). 
For the regeneration of cytosolic ATP, the newly formed ATP can be exported from the 
mitochondria via the ATP/ADP carrier protein. This process, however, requires the 
operation of the phosphate carrier Mit1p, which translocates phosphate together with a 
proton from the cytosol to the mitochondria.

**Cell energetics of compartmented networks**

The lack of the complex I in the respiratory chain in *S. cerevisiae* has direct consequences 
on the energetics of its metabolic network. Unlike in most other eukaryotic cells, cytosolic 
and mitochondrial NADH are equivalent in terms of their potential energy yield. Shuttling 
of the NAD/NADH couple between both compartments has no consequences for the 
energy content, as it would for e.g. *C. utilis*. In a normal eukaryote with a functional 
complex I, energy-independent shuttling of redox equivalences from the mitochondria to 
the cytosol would thus result in an loss of approximately one ATP for each NADH. 
Therefore, a shuttle like the ethanol/acetaldehyde system would, if existent, be tightly 
regulated to prevent unnecessary waste of energy. In contrast, *S. cerevisiae* can allow the 
free operation of this system.

A further aspect of this difference can be found in the TCA cycle. Since the localization 
of the dehydrogenases will not influence the energy yield, the transport of its metabolites 
can be optimized to minimize transmembrane trafficking while allowing the extensive use 
of cytosolic dehydrogenases. Cells with a functional complex I will be more likely to 
restrict their dehydrogenases to the mitochondria, hence producing more of the NADH 
inside the mitochondria, and thus increasing the flux through complex I and maximizing 
the energy yield. In the this case, the anabolic requirements for the C₃-intermediates will 
be more likely to be fulfilled by direct export of oxoglutarate, rather than the export of 
citrate and its cytosolic conversion to the first. Similarly, import and mitochondrial 
oxidation of the fumarate will be energetically more favourable in these cells, than a 
cytosolic pathway.

Although most reactions of the central metabolism are known, many processes e.g. protein 
folding, protein turnover, intracellular transport processes and maintenance, and are poorly 
understood in terms of its energy requirement. For modelling purposes these unknowns are 
commonly lumped together in the *P/O* ratio and the maintenance coefficient (van Gulik
and Heijnen, 1995; Vanrolleghem and Heijnen, 1998). The $P/O$ ratio is the amount of ATP generated per electron pair and thus summarily describes the efficiency of the energy transfer from NADH or FADH$^+$ to ATP in the respiratory chain and the F$_{0}$F$_{1}$-ATPase. $P/O$ values obtained with un compartmented models range from 1.1 to 1.4 based on measured yields for yeast grown in chemostat cultures (Stückrath et al., 2001; van Gulik and Heijnen, 1995; Vanrolleghem and Heijnen, 1998; Verduyn et al., 1991). The maintenance $k'_{ATP}$, often divided into a growth dependent and a growth independent part, summarizes the energetic requirements, which were not taken into account by the metabolic network. Reported values for the $k'_{ATP}$ lie around 800 mmol/C-mol biomass for yeast grown at a dilution range of 0.1 hr$^{-1}$.

These values can not be directly compared with a compartmented metabolic network model. Here, a certain number of protons should be allowed in a model to pass through the inner mitochondrial membrane without generating ATP as observed in the brown fat uncoupling protein (UCP) to allow for a certain inefficiency in the energy transfer process. Compared with an un compartmented model, a model with intracellular transport can theoretically distinguish between the efficiency of the respiratory chain and the efficiency of the ATPase. Mathematically, this can be accomplished by modelling the respiratory chain and the ATPase with a variable proton translocation stoichiometry. However, mathematically it is more convenient to add an additional reaction, termed proton slip, to the network, which allows the free travel of protons across the mitochondrial membrane. The efficiencies of the respiratory chain and the ATPase can then be determined experimentally by the comparison of the flux ratios of the proton slip, respiratory chain, and ATPase under different growth conditions.

Proposed metabolic networks

Based on the presented evidence, we attempted to build consistent models for the growth of yeast on either glucose, ethanol, or acetate with ammonia as carbon source (figure 7).

The C$_{1}$-metabolism was assigned solely to the cytosol, as was the associated glycine formation, despite the additional presence of mitochondrial enzymes due to the fact, that the demand of C$_{1}$-carbon units is located exclusively in the cytosol for the reactions included in our model. The GOGAT system is assumed to be cytosolic, due to the absence of a phenotype for the deletion mutant. Thus, all ammonia fixation is located in the cytosol and a transport of an ammonia group carrier is required. Of the proposed shuttle systems, the shuttle using o xoisovalerate is more likely than one involving oxaloacetate, not least because both required transporters are also needed operating in the same direction during the synthesis of the branch chain amino acids. Hence, a valine/oxoisovalerate shuttle was chosen for the model (figure 3 F).

The pyruvate branch point was modelled with a NADP dependent acetaldehyde dehydrogenase as standard pathway. For growth on ethanol an additional NAD dependent one has to be added to enable the adjustment of the ratio of NADPH to NADH produced.
Figure 7: compartmented metabolic network used for modelling, cytosolic reactions are in black, mitochondrial ones are in grey, small numbers indicate the amount of molecules consumed, if larger than two

according to the needs of the anabolism. The mitochondrial and the cytosolic alcohol dehydrogenases are equivalent from a modelling standpoint, as long as an alternative shuttle mechanism for the redox potential is provided within the model. The cytosolic one was subsequently chosen. The malic enzyme was not included in the model for neither
growth on ethanol nor acetate, since it is energetically almost undistinguishable from the alternative pyruvate provision via the PEP carboxykinase.

Four systems were considered for the import of the C₂-body into the mitochondrial TCA cycle. While strong evidence was found against the operation of a carnitine shuttle, all three other options can not easily be discarded. As stated before, these possibilities require different amounts of energy for the various transport steps. A metabolic model might be able to distinguish between the dual location of the acetyl-CoA synthases, the cytosolic location of the citrate synthase, or an acetyl-CoA shuttle around the isopropyl-malate synthase. All three version should therefore be initially implemented and tested in a metabolic network as possible alternatives.

While the TCA cycle is generally modelled as being mitochondrial, the regeneration of aspartate from fumarate was implemented in the cytosol. This minimises the amount of intracellular metabolite trafficking while not influencing the energy yield in yeast. Antiporters were used whenever possible for the transport of TCA cycle intermediates. Thus, during growth on glucose, the model used a oxoglutarate/malate shuttle encoded by ODC1&2, growth on C₂-compounds was modelled with additional succinate/fumarate and (iso)citrate/malate shuttles to minimize the energetic costs of trafficking. In all cases, the oxaloacetate transporter is used to balance the out-flux of mitochondrial TCA-cycle compounds for anabolic purposes.

Alanine production was chosen to be cytosolic, however a mitochondrial localisation would also be justified. The first step of the lysine synthesis pathway was modelled as cytosolic, the ones form homocitrate to oxoadipate as mitochondrial, and the remaining again as cytosolic. Homocitrate was imported via active transport into the mitochondria, oxoadipate was exchanged via the ODC1 carrier against malate. Leucine formation was placed in the cytosol. The implemented transporters are indicated through an circle around the metabolite in figure 7.

The model’s use of an internal and an external NADH dehydrogenase provided for a redox shuttle. This annulled the requirements for further NADH shuttle systems, e.g. involving the ones depicted in figure 3.

Despite the marked differences, modelling of ethanol and acetate growth was successfully done using the same pathways for both substrates together with one set of energetic parameters (Stückrath et al., 2001). This is further supported by the results of Verduyn et al. who calculated the same P/O ratio for ethanol and acetate using the same biomass composition in both cases (Verduyn et al., 1991). We will therefore use two basically identical networks for the two substrates, despite the different phenotypes observed.

Conclusions

The localisation of metabolic pathways is not fully understood at present, although an abundance of localization studies have been performed for proteins in S. cerevisiae.
Generally, insufficient emphasis had been laid on the variations due to different growth conditions. Nevertheless, a general picture emerges, which allows the placement of most pathways in one of the compartments of the cell and their linkage through appropriate transporters.

While most reactions occur in the cytosol, some anabolic pathways for the synthesis of amino acids are confined to the mitochondria. The TCA cycle and the respiratory chain, thus the reactions of central energy metabolism, are distributed over both compartments and function only in combination with closely regulated transport processes across the inner mitochondrial membrane. Here, the compartmentation is of utmost importance for the functioning of the metabolism, thus models which do not include a mitochondrial compartment are likely to fail when applied to the validation of the cell’s energetics. In contrast, the peroxisomes are the foremost location of the β-oxidation, but of little significance during growth on substrates other than long chain fatty acids. Thus, restriction of a model to two intracellular compartments should not significantly influence its predictive capacities for cells grown on glucose, ethanol, or acetate.

Information on the existence of intracellular transport proteins is still sparse. The understanding of the energy requirements of the identified transport steps is only at its beginning. The setup of a functional model thus still requires a substantial number of assumptions. However the current knowledge on the transport systems in yeast allows to draw some interesting conclusions on the energetic consequences of the compartment, as will be shown in a related paper (Lange and Heijnen, 2002).

The presented work presents the current state of knowledge on the compartmental localisation of the metabolic pathways in Saccharomyces cerevisiae. Several important questions remain yet unsolved, especially around the mitochondrial import of \( C_2 \)-bodies. Here, different possibilities were raised, among them a new hypothesis for an acetyl-CoA shuttle using the isopropyl malate synthase. A similar shuttle system was proposed for the amino group import using the branch chain amino acid transferases (BAT1 & BAT2). A compartmented model is thus a valuable tool for expanding the understanding of the metabolism in yeast. A scheme of such a network is given in figure 7, which provides a basic compartmented metabolic network for the central metabolism.

Acknowledgements

I would like to thank Prof. Jack Pronk for his critical remarks and valuable suggestions, Dr. Preben Krabben for the stimulating discussions and his contributions, especially during the elucidation of the respiratory chain. The work was financed through the European framework IV program on cell factories and is part of the project ‘From gene to product in yeast: a quantitative approach’.
References


93


Chapter 5

Energetic aspects of the compartmentation of the metabolism

submitted as

Energetic aspects of the compartmentation of the metabolism in *Saccharomyces cerevisiae*

H.C. Lange, J.J. Heijnen

to

Biotechnology and Bioengineering
Energetic aspects of the compartmentation of the metabolism

Abstract

Summarizing the current knowledge on enzyme localisation, a strictly compartmented model of the metabolism was developed and used for the determination of metabolic fluxes, including fluxes across the mitochondrial membrane, in aerobically grown S. cerevisiae. The compartmentation of the respiratory chain and the $F_0F_1$-ATPase with their related transport processes was analysed for its impact on the energy metabolism and found to cause a higher energetic cost for the regeneration of cytosolic ATP compared to mitochondrial ATP. The introduction of a separate mitochondrial compartment and the inclusion of the appropriate transport reactions decreased the amount of cell’s Gibbs energy consumption, which is not accounted for by the metabolic network, by about 30%. Based on the experimental data from chemostat cultivations, a new set of four energetic parameters were deduced. Two different metabolic networks were compared and resulted in $P/O$ ratios of 1.30 and 1.18, a $k_{ATP}$ of 300 and 170 mmol/Cmol, and a $m_{ATP}$ of 49 and 43 mmol/Cmol/hr.
Introduction

The cell metabolism of eukaryotic cells is distributed between several organelles. Of these, the mitochondrion has a special role under aerobic conditions as the primary energy provider, regenerating the main energy carrying compound ATP at its inner membrane. The compartmented organisation of the cell has thus an crucial influence on the energy transformation via the respiratory chain and the F₉F₁-ATPase, but also on its consumption due to various energy requiring transport processes involved in the energy metabolism and the provision of anabolic precursors. In the presented work, we attempt to model the cell metabolism of *S. cerevisiae* with a strict, two-compartmental model, which does not only include the metabolic reactions, but also the required transport steps.

The presented compartmented model describes the cell’s metabolism in a more realistic and detailed form than previous uncompartmented models (Stückrath et al., 2001; van Gulik and Heijnen, 1995; Vanrolleghem et al., 1996). Besides providing a better basis for metabolic flux analysis in combination with the isotope labelling approach (Christensen and Nielsen, 1999), it is also capable of providing a predictive capacity for metabolic situations, which are only comprehensible in the context of a compartmented environment, such as the pyruvate decarboxylate deletion mutant (*pdc1Δ pdc5Δ pdc6Δ*), which requires an additional acetyl-CoA source for growth on glucose despite the presence of a functional pyruvate dehydrogenase - the former is required in the cytosol, the latter provides in the mitochondria (Flikweert et al., 1999).

Compared with earlier works using compartmentation in yeast (Granström et al., 2000; Nissen et al., 1997)(Ostergaard et al., 2001; Wahlbom et al., 2001) the presented model of aerobic growth strictly separates the different pools of metabolites and only allows the transport into a different compartment via a designated transport reaction. Transporters are modelled according to current literature, thus also including the energy requirements for the translocation. This has direct implications for the energetic value for e.g. ATP depending on its localisation, which was ignored in the previous publications. It further allows the estimate of the energy burden due to transport, which to our knowledge has not be reported elsewhere.

Yield predictions of mutants based on metabolic network analysis as shown by Stückrath et al. demand a solid understanding of the cell’s energetics (Stückrath et al., 2001). Previous works have demonstrated the successful application of a three-parameter model for aerobic growth, in which the P/O ratio describes the efficiency of the ATP generation from NADH, \(k_{\text{ATP}}\) summarizes the unknown growth related ATP requirements, and \(m_{\text{ATP}}\) accounts for the time independent ATP consumption for maintenance (van Gulik and Heijnen, 1995; Vanrolleghem et al., 1996; Verduyn, 1992). We expanded our model for an compartmented network, splitting the P/O ratio into two terms, describing the efficiency of the respiratory chain and the F₉F₁-ATPase. This allows the application of this approach on any compartmented model, independently of the configuration of the respiratory chain. This also lead to a new theoretical maximum for the P/O ratio of 1.38 for compartmented networks in *S. cerevisiae*.
The compartmented model is further able to categorize the different transport processes through the determination of the transmembrane fluxes under the investigated growth conditions. Calculations of the associated energetic costs of metabolite trafficking is thus possible. Based on these calculations, we were able to discuss the physiological significance of the succinate/fumarate and (iso)citrate/malate antiports for cells grown on ethanol or acetate.

Compartmentation of the metabolism in yeast

A model of the yeast metabolism was set up with two compartments, cytosol and mitochondria, for the biomass formation from glucose, ethanol, and acetate as C-source and ammonia as N-source. Metabolite translocations over the cell- and the mitochondrial membrane were limited to specific transport reactions, thus creating distinct pools for all metabolites. The mitochondria contained the reactions of the TCA cycle, the synthesis reactions of the amino acids valine, isoleucine, and ornithine, as well as parts of the lysine synthesis. The TCA cycle reactions aconitase, malate dehydrogenase, and fumarate hydratase as well as the branched chain amino acid transferase for valine, and the reactions from acetaldehyde to acetyl-CoA were placed in both compartments, all other reactions were strictly cytosolic. The branch chain amino acid transferase, encoded by BAT1 and BAT2, was used for the shuttling of the amino group donor couple, glutamate and oxoglutarate.

Five distinct networks were used for the modelling of growth of *S. cerevisiae* of the three substrates glucose, ethanol and acetate, differing around the pyruvate branchpoint (figure 1). Glucose metabolism was modelled with the glycolysis and a parallel pentose-phosphate (PP) pathway, the latter supplying the cytosolic NADPH. In this case, pyruvate as central metabolite provided both the cytosolic and the mitochondrial acetyl-CoA via pyruvate decarboxylase and pyruvate dehydrogenase respectively. Anaplerotic requirements of the TCA cycle were met via the pyruvate carboxylase. For growth on the C3-compounds ethanol and acetate, the glyoxylate cycle reactions isocitrate lyase and malate synthase provided the necessary precursors for the anabolism. Phosphoenol-pyruvate (PEP) carboxykinase supplied for these substrates the starting material for the gluconeogenesis; the glucose-6-phosphate dehydrogenase was absent in these networks. Instead, the cytosolic NADPH was generated via the NADP-dependent acetaldehyde dehydrogenase for ethanol, and via a cytosolic, NADP-dependent isocitrate dehydrogenase for acetate.

The compartmental placement of the acetyl-CoA synthases (Acs1p & Acs2p) is still subject to discussion (Klein and Jahnke, 1968) and recent publications indicate a cytosolic localization (de Jong-Gubbels, 1998; van den Berg and Steensma, 1995). This would, however, conflict with the mitochondrial (Cit1p, Cit3p) and peroxisomal (Cit2p) localization of the citrate synthase and the absence of an operating acetyl-CoA transport mechanism. Therefore, two different scenarios were modelled, differing in the localization of the acetyl-CoA synthase and the citrate synthase. In the network I a strictly cytosolic occurrence was used, while network II included an additional mitochondrial reaction (figure 1).
Figure 1: central metabolism for the substrates glucose (G), ethanol (E-I and E-II), and acetate (A-I and A-II) using network I and II respectively. For malate, the direction of the arrow shows the net-direction of the combination of both antiports.
Intracellular transporters

The distribution of metabolites across two compartments requires the trafficking via specific transporter proteins across the mitochondrial membrane. In the following these intracellular transporters are briefly discussed for the anabolic pathways, the catabolic pathways, the energy metabolism, and the TCA cycle.

Little is known about the transporters in anabolism. Exceptions are a.) the o xoadipate antiport (Odc1p & Odc2p) in the lysine synthesis pathway, which functions as an antiport using any oxocarboxylate as substrate (Palmieri et al., 2001) and b.) the ornithine transporter (Ort1p), a proton antiport, thus able to actively export the mitochondrially synthesised ornithine. For all other precursors (oxoisovalerate, valine, homocitrate, threonine, oxobutyrate, isoleucine, ammonia) an energy dependent transport was assumed; import of the compound into the mitochondria was modelled with a proton symport, export via a proton antiport, thus using in both cases the proton-motive force.

Carrier proteins in catabolic pathways have not yet been identified. Ethanol and acetaldehyde are supposed to diffuse freely across the mitochondrial membranes. Growth on glucose requires a substantial flux of pyruvate into the mitochondria in addition to the anabolic requirements of the oxoisovalerate synthesis. However, no transport protein for pyruvate has yet been found. Therefore, two different versions were tested: proton symport together with pyruvic acid (PYR) and with pyruvate (PYR).

Carriers involved in the energy metabolism are generally better characterized. The proteins of the respiratory chain translocate six protons from the mitochondria to the cytosol for each electron pair received. Due to the lack of a functional complex I in yeast (Bakker et al., 2001), mitochondrial and cytosolic NADH as well as FADH₂ are energetically equivalent. Ten translocated protons are thought to be channelled back into the mitochondria to regenerate three ATPs via the FₒF₁-ATPase complex of S. cerevisiae (Stock et al., 1999), a ratio recently also found in E. coli (Jiang et al., 2001). The non-integral number of protons per ATP is still subject of discussions, but increasingly accepted (Arechaga and Jones, 2001). The generated ATP can be transported into the cytosol via two different ATP³⁻/ADP²⁻ exchange carrier proteins (Aac1p & Aac2p). This requires the operation of the phosphate carrier protein encoded by MIR1, a H₃PO₄/H⁺ symport, as importer. Thus, the regeneration of one mol cytosolic ATP from mitochondrial ATP results in the net-import of one proton into the mitochondria (Reich and Selkov, 1981).

Biomass formation requires the net transport across the mitochondrial membrane of the TCA cycle intermediates oxaloacetate and oxoglutarate to provide precursors for anabolic reactions. A mitochondrial oxaloacetate transporter (OAC1) was characterized earlier as proton symport (Palmieri et al., 1999a). Alternatively and energetically equivalent, the dicarboxylate transport protein Dic1p is able to translocate malate across the membrane in exchange for phosphate (Palmieri et al., 1999b). The subsequent malate-oxaloacetate conversion is possible in both compartments. Both systems are energetically equal due to
the equivalence of cytosolic and mitochondrial NADH; the direct proton dependent oxaloacetate carrier was initially considered in the model and will be further discussed in the result section. The apparent absence of an oxoglutarate carrier in the yeast genome (Palmieri et al., 1999a) can be overcome via one of the two oxodicarboxylate carriers (Odc1p or Odc2p) which facilitate a malate/oxoglutarate exchange across the mitochondrial membrane (Palmieri et al., 2001). The imported malate can subsequently be re-exported via the previously discussed oxaloacetate carrier system (figure 1).

**Figure 2:** main processes involved in maintaining the proton balance across the mitochondrial membrane: export of protons through the respiratory chain; import via the FoF1-ATPase, active transport of metabolites, transport of phosphate.

Activity of the glyoxylate cycle during growth on C2-carbon sources requires the mitochondrial import of cytosolically synthesized succinate via the succinate-fumarate antiport (SFCI) (Palmieri et al., 1997). The fumarate can be reimported into the mitochondria either via the oxaloacetate carrier, after cytosolic hydration and oxidation, or via electroneutral exchange of (iso)citrate and malate through the tricarboxylate carrier protein Ctp1p (Kaplan et al., 1995; Palmieri et al., 1972). The latter system exports the precise amount of (iso)citrate required by the cytosolic isocitrate lyase, thus allowing the bi-compartmental glyoxylate cycle to operate without energy requirements for transmembrane trafficking. The last system was therefore chosen for the model, the resulting energy savings will be quantified later (figure 1 E&A). As a consequence of the dual substrate specificity of the tricarboxylate carrier protein (Ctp1) the cytosolic localisation of the aconitase is not strictly required and can be compensated for by a citrate/malate exchange, a isocitrate/malate exchange, or a combination of both.
The requirement of a mitochondrial source of an ammonia group was solved with the introduction of an oxoisovalerate-valine shuttle system, which regenerates mitochondrial glutamate via the deamination of valine to oxoisovalerate.

**Energetics of compartmented metabolic networks**

The energy demand of the cell was modelled as an extension of previous works (van Gulik and Heijnen, 1995; Vanrolleghem et al., 1996) with four parameters, which describe the efficiency of energy conversion during ATP formation from H\(^+\) import (P/O and \(\alpha\)), the unaccounted energy demand of the biomass formation \(k_{ATP}\) and the maintenance requirement \(m_{ATP}\) of a cell independently of its growth rate. ATP regeneration involves four principle reactions:

\[
\begin{align*}
q_{2e^-} : & \quad 2e^- + 2H^+ + 0.5O_2 + 6 \cdot \alpha_1 \cdot H^+_{\text{mit}} \rightarrow H_2O + 6 \cdot \alpha_1 \cdot H^+_{\text{cyt}} \\
q_{ATP} : & \quad ADP_{\text{mit}} + Pi_{\text{mit}} + 3\frac{1}{2}/\alpha_2 \cdot H^+_{\text{cyt}} \rightarrow ATP_{\text{mit}} + H_2O + 3\frac{1}{2}/\alpha_2 \cdot H^+_{\text{mit}} \\
q_{ATP} : & \quad ATP_{\text{mit}} + ADP_{\text{cyt}} \rightarrow ATP_{\text{cyt}} + ADP_{\text{mit}} \\
q_{Pi} : & \quad Pi_{\text{cyt}} + H^+_{\text{cyt}} \rightarrow Pi_{\text{mit}} + H^+_{\text{mit}}
\end{align*}
\]

(1)

Due to the absence of complex I in the respiratory chain of *S. cerevisiae*, the flux of protons translocated by \(q_{2e^-}\) is independent of the source of the electrons, be it from cytosolic or mitochondrial NADH or from FADH\(_2\). The efficiencies of the respiratory chain and the F\(_0\)F\(_1\)-ATPase are described with the factors \(\alpha_1\) and \(\alpha_2\) respectively, both lying between zero and one. \(\alpha_1\) relates to the ratio of protons exported in the complexes I-IV per electron pair, while \(\alpha_2\) relates to the amount of protons required for the generation of one ATP.

The energetic parameters can be obtained by solving the balances for ATP, protons and phosphate around the mitochondrial membrane (figure 2). The mitochondrial proton balance is written as

\[
6 \cdot \alpha_1 \cdot q_{2e^-} = 3\frac{1}{2}/\alpha_2 \cdot q_{ATP} + q_{H^+} + q_{Pi} \tag{2}
\]

where \(q_{2e^-}\) denotes the flux of pairs of electrons through the respiratory chain in mol/C-mol biomass/hr. The flux \(q_{H^+}\) denominates the net proton flow from the cytosol to mitochondria associated with active transport processes excluding the proton flux \(q_{Pi}\) associated with the symport of phosphate. Finally, the mitochondrial proton influx via the F\(_0\)F\(_1\)-ATPase is described as a multiple of \(q_{ATP}\).

The two equations representing the mitochondrial and cytosolic ATP balances are:

\[
\begin{align*}
\text{mit.:} & \quad q_{ATP} + \sum q_{ATP-network}^{\text{mit}} - q_{ATP} = k_{ATP}^{\text{mit}} \cdot \mu + m_{ATP}^{\text{mit}} \\
\text{cyt.:} & \quad q_{ATP} + \sum q_{ATP-network}^{\text{cyt}} = k_{ATP}^{\text{cyt}} \cdot \mu + m_{ATP}^{\text{cyt}}
\end{align*}
\]

(3)

The net ATP production of the various reactions in the network is summarized in \(\sum q_{ATP-network}^{\text{cyt}}\) and \(\sum q_{ATP-network}^{\text{mit}}\) for the cytosol and the mitochondria respectively. The flux \(q_{ATP}^{\text{trans}}\) quantifies the mitochondrial ATP export via the ATP/ADP carrier. ATP
consumption not related to metabolic network reactions is accounted for by a separate \( k_{ATP} \) and \( m_{ATP} \) for each compartment.

No phosphorylated compounds are transported across the mitochondrial membrane. Neglecting the accumulation of phosphate-containing mitochondria due to biomass formation, one can write the phosphate balance as:

\[
q_{Pi-trans} = q_{ATP-trans}
\]  

(4)

Combining these balances (equations 2, 3, and 4), the energetic parameters can be related with

\[
6\alpha_1 \cdot q_{2e} + \left(3\frac{1}{3}/\alpha_2 + 1\right)\left(\sum q_{ATP-network}^{\text{c}} + k_{ATP}^{\text{c}} \cdot \mu - m_{ATP}^{\text{c}}\right) + 3\frac{1}{3}/\alpha_2 \left(\sum q_{ATP-network}^{\text{mit}} - k_{ATP}^{\text{mit}} \cdot \mu - m_{ATP}^{\text{mit}}\right) - q_{H^+-trans} = 0
\]  

(5)

to the four rates \( q_{2e} \), \( q_{ATP-network}^{\text{c}} \), \( q_{ATP-network}^{\text{mit}} \), and \( q_{H^+-trans} \) which can be determined by solving the flux balance of the metabolic network.

The P/O ratio is defined as the amount of cytosolic ATP produced per electron pair passed through the respiratory chain to enable direct comparison with an uncompartmented model. Thus, with

\[
P / O_{\text{compartmented}} = \frac{6\alpha_1 \cdot \alpha_2}{3\frac{1}{3} + \alpha_2}
\]  

(6)

and

\[
\alpha := \frac{3\frac{1}{3}}{3\frac{1}{3} + \alpha_2}
\]  

(7)

equation (5) simplifies to

\[
P / O_{\text{comp}} \cdot q_{2e} \left(\sum q_{ATP-network}^{\text{c}} + \alpha \cdot \sum q_{ATP-network}^{\text{mit}}\right) - k_{ATP}^{\text{c}} \cdot \mu - m_{ATP}^{\text{c}} + \alpha \cdot m_{ATP}^{\text{mit}} - (1 - \alpha) \cdot q_{H^+-trans} = 0
\]  

(8)

As \( \alpha_2 \) is defined as lying between zero and one, \( \alpha \) can assume only values between 0.77 and one. It can be expected that \( \alpha_2 \) is close to unity, and therefore \( \alpha \) close to 0.77. It is thus evident, that under physiological conditions the regeneration of the cytosolic ATP is costlier in terms of protons being translocated back into the mitochondria than the regeneration of mitochondrial ATP. This is due to the transport cost for the phosphate group, which is required in the mitochondria for ATP formation from ADP. It also becomes obvious, that \( k_{ATP}^{\text{c}} \) and \( k_{ATP}^{\text{mit}} \) cannot be determined independently and should be combined in a general \( k_{ATP} \) as should be both maintenance coefficients in \( m_{ATP} \). This gives:

\[
P / O_{\text{comp}} \cdot q_{2e} + \left(\sum q_{ATP-network}^{\text{c}} + \alpha \cdot \sum q_{ATP-network}^{\text{mit}}\right) - k_{ATP} \cdot \mu - m_{ATP} - (1 - \alpha) \cdot q_{H^+-trans} = 0
\]  

(9)
Energetic aspects of the compartmentation of the metabolism

where all fluxes $q_i$ can be determined by metabolic flux analysis without \textit{a priori} assumptions about the four energetic parameters. Thus, the energetic parameters can be derived from a minimum of four independent measurements sets, analogous to the ones of an uncompartmented model (Vanrolleghem and Heijnen, 1998).

The maximal theoretical $P/O$ value for a compartmented network calculates to 1.38 with both the respiratory chain and the ATPase functioning with an efficiency of 100%, thus $\alpha_1 = 1$ and $\alpha_2 = 1$. If the mitochondrial phosphate transport is neglected or absent (as e.g. in prokaryotes) the uncompartmented $P/O$ ratio becomes

$$P/O_{\text{uncompartmented}} = \frac{6}{3^{\frac{1}{3}}} \cdot \alpha_1 \cdot \alpha_2$$

with a maximal theoretical value of 1.80.

Results

The original data set was obtained from fourteen carbon limited chemostat cultivations of \textit{S. cerevisiae} grown at dilution rates from 0.02 to 0.21 hr$^{-1}$ on glucose, ethanol, and acetate as reported earlier (Stückrath et al., 2001). The number of different measurements exceeded the required minimum for a complete flux determination. Therefore, the additional information was used together with the elemental balances included in the metabolic networks for statistical evaluation and data reconciliation according to van der Heijden et al. (van der Heijden et al., 1994a; van der Heijden et al., 1994b). All data sets were statistically acceptable, the reconciled data of the main exchange fluxes is presented in table I.

<table>
<thead>
<tr>
<th>steady state cultivations of \textit{S. cerevisiae}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>substrate</strong></td>
</tr>
<tr>
<td>ethanol</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

\textbf{Table I:} main reconciled exchange fluxes per C-mol biomass obtained from chemostat fermentations of \textit{S. cerevisiae}
Based on the reconciled data, the metabolic fluxes for the different situations were calculated. Most fluxes were found to proceed in their expected directions, however, the model predicted a negative flux of 200 mmol/C-mol biomass/hr through the cytosolic malate dehydrogenase for glucose grown cultures, while on C₂-carbon sources the flux proceeded from malate to oxaloacetate. Since the reaction catalysed by malate dehydrogenase is reversible, this was deemed admissible.

**energetic parameters**

<table>
<thead>
<tr>
<th></th>
<th>model employed</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>uncomp.</td>
<td>network I</td>
<td>network II</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>n.a.</td>
<td>0.94</td>
<td>0.85</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>n.a.</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$P/O_{cyt}$</td>
<td>1.44</td>
<td>1.30</td>
<td>1.18</td>
</tr>
<tr>
<td>$k_{ATP}$</td>
<td>646</td>
<td>300</td>
<td>170</td>
</tr>
<tr>
<td>$m_{ATP}$</td>
<td>55</td>
<td>49</td>
<td>43</td>
</tr>
<tr>
<td>$P/O_{cyt}^{theor}$</td>
<td>1.80</td>
<td>1.38</td>
<td>1.38</td>
</tr>
</tbody>
</table>

**Table II:** derived energetic parameters based solely on data of glucose and ethanol cultures; and, for comparison, theoretically possible maximum P/O ratio expressed in terms of cytosolic ATP

Of the transport reactions, all proceeded in the expected direction with the exception of the initially used oxaloacetate proton symport. An outward flux of 2 mmol oxaloacetate/C-mol biomass was observed for all cases, due to the operation of the Odc2p-antiport used to export the oxoacidate from the mitochondria. The resulting import of oxoglutarate was not balanced by the use of mitochondrial precursors of the TCA cycle, e.g. for the mitochondrial formation of ornithine. The oxaloacetate/proton symport (Oac1p) was therefore not used in the model and replaced with an oxaloacetate exporter operating as a proton antiport. This is consistent with lack of a phenotype of an oac1Δ mutant on glucose.

**Energetic parameters**

Using the reconciled fluxes, the energetic parameters for both networks are calculated using equation (9) based on the ethanol and glucose limited chemostat cultures as outlined earlier. The data sets required setting upper boundary conditions for the efficiencies $\alpha_1$ and $\alpha_2$ during the optimum search. A sensitivity analysis revealed a very high influence of possible measuring errors on $\alpha_1$ and $\alpha_2$, while the combined parameter P/O was very robust. The solution of the linear optimization problem consistently resulted in an $\alpha_2$ at the upper boundary condition of one. The obtained energetic parameters are stated in table II for both networks: Network I resulted in a P/O ratio of 1.30, close to its theoretical maximum, demanding a combined efficiency of 94% for the respiratory chain and the
Energetic aspects of the compartmentation of the metabolism

$F_{0}F_{1}$-ATPase. The lower $P/O$ ratio of 1.18 in network II is compensated for by a lower $k_{ATP}$ (table II).

As a test, the different parameter sets are used together with their appropriate network to predict the biomass yield on acetate, for which the data was not used in the determination of the four parameters. The results compare excellently with the actual measurement, thus giving an independent verification of the method and the obtained parameters (table III). Network II performed marginally better than network I without the mitochondrial acetyl-CoA synthase.

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Measured</th>
<th>Uncomp.</th>
<th>Compartmented Network I</th>
<th>Compartmented Network II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>0.598</td>
<td>0.600</td>
<td>0.600</td>
<td>0.604</td>
</tr>
<tr>
<td>Ethanol</td>
<td>0.623</td>
<td>0.612</td>
<td>0.617</td>
<td>0.614</td>
</tr>
<tr>
<td>Acetate</td>
<td>0.382</td>
<td>0.380</td>
<td>0.365</td>
<td>0.384</td>
</tr>
</tbody>
</table>

Table III: yield predictions for compartmented models I and II based on energetic parameters of Table II and for un compart mented model from earlier works (Stückrath et al., 2001) for a growth rate 0.1 hr$^{-1}$

Comparison with the un compartmented model shows, that the growth related energy requirement $k_{ATP}$ decreases more than 50% for the compartmented networks, which can be explained by the inclusion of the energy requirements for transport. The cytosolic ATP based compartmented $P/O$ ratio is relatively closer (94% and 84%) to the theoretical maximum, compared to the estimated un compartmented P/O ratio of 1.44, which represents an 80% efficiency. The maintenance $m_{ATP}$ decreases slightly for the compartmented models, the difference being more pronounced for network II.

The influence of the choice of transport mechanism was tested on the example of pyruvate transport. The differences in the energetic parameters between an electro-neutral and an active transport with proton symport was less than 1% for the P/O ratio and the maintenance coefficient $m_{ATP}$. The growth related ATP requirement $k_{ATP}$ increased by 60 mmol/C-mol for the energetically less expensive electro-neutral carrier system.

Fluxes related to energy metabolism

The proton flux across the mitochondrial membrane encompasses the majority of the cell's energy household and is summarized in equation (2). Energy is delivered to the respiratory chain in the form of 130 mmol of reduced adenine dinucleotides/hr per C-mol glucose grown biomass at a growth rate of $\mu = 0.1$ hr$^{-1}$. At the same $\mu$, biomass built from C$_2$-carbon sources provides a higher flow: 283 and 328 mmol/C-mol/hr for ethanol and acetate respectively. This translates in an outward flux of protons between 649 and 1972 mmol/C-mol/hr (table IV), with C$_2$-grown cells exporting about 2.5-times this amount compared to glucose grown ones. The various terms of the proton balance are given in
The smaller proton flux in network II compared to network I is due to its smaller $\alpha_i$.

The generated proton motive force is used in the $F_0F_1$-ATPase and for active transport processes. At a growth rate of 0.1 hr$^{-1}$ around 600 and 1400 mmol protons/C-mol/hr are used to regenerate the ATP consumed during biomass formation from $C_6$ and $C_2$ carbons respectively (table IV). This represents around 80% of the energy provided by the respiratory chain. The remaining 13 to 20% of the translocated protons is be used for the intracellular trafficking across the mitochondrial membrane. Here, the lion’s share is consumed for the phosphate translocation, accounting for 75% in glucose grown cells and up to 97% in cells grown on $C_2$-carbon sources. In network I 326 mmol protons/C-mol biomass/hr are consumed for transport, thus substantially more than the 226 mmol/C-mol/hr required in network II when grown on ethanol (447 vs. 226 mmol/C-mol/hr for acetate). In these cases, about one fifth of the total available energy is used in transport related to regeneration of cytosolic ATP, which is mainly due to its high demand in the activation step for acetate (table IV).

**Fluxes in the energy metabolism** (mmol H$^+$/C-mol/hr)

<table>
<thead>
<tr>
<th>network</th>
<th>$6 \times \alpha_1 \times q \times 2e^-$</th>
<th>$10/3 \times \alpha_2 \times q \times ATPase$</th>
<th>$q \text{ H}^+ \text{-transport}$</th>
<th>$q \text{ Pi-transport}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethanol</td>
<td>675</td>
<td>1349</td>
<td>12</td>
<td>314</td>
</tr>
<tr>
<td>II</td>
<td>1483</td>
<td>1259</td>
<td>12</td>
<td>214</td>
</tr>
<tr>
<td>acetate</td>
<td>2005</td>
<td>1558</td>
<td>13</td>
<td>435</td>
</tr>
<tr>
<td>II</td>
<td>1656</td>
<td>1431</td>
<td>12</td>
<td>214</td>
</tr>
<tr>
<td>glucose</td>
<td>740</td>
<td>617</td>
<td>31</td>
<td>93</td>
</tr>
<tr>
<td>II</td>
<td>655</td>
<td>532</td>
<td>30</td>
<td>93</td>
</tr>
</tbody>
</table>

Table IV: calculated fluxes of reactions involved in the energy metabolism of the inner mitochondrial membrane expressed in terms of protons (equation 2) at a growth rate of $\mu = 0.1$ hr$^{-1}$

The distribution of proton fluxes in and out of the mitochondria is depicted in figure 3 for glucose grown cells over a range of different dilution rates. Network I sustains a 12% higher proton flux for all growth rates. The difference is mostly due to the higher energy demands for maintenance and growth related ATP requirements (dashed line). The sum of both account for 41 and 47% of the ATP consumption at a growth rate of 0.1 hr$^{-1}$ in networks I and II, respectively. This is substantially less than the 65% which are unaccounted for by the metabolic network in an uncompartmented model, where 123 out of 190 mmol ATP/C-mol/hr are used for maintenance and unspecified, growth related metabolic activities.

Summarizing, the share of energy in form of a potential proton-motive force, which is not accounted for by the metabolic network and which is presumably used for maintenance, metabolic activities not included in the metabolic network, or lost due to inefficiencies, drops from 30 to 22% for ethanol, and from 26 to 19% for acetate grown cells for network
I with a cytosolic acetyl-CoA synthase, and to, respectively, 19% and 16% for network II with an additional mitochondrial localization of the acetyl-CoA synthase.

![Graph](image)

**Figure 3:** proton export (filled symbols) from and import (open symbols) into the mitochondria for glucose grown cultures of *S. cerevisiae*, via respiratory chain $q_{II:respiratory\ chain}$ (■), ATPase $q_{II:ATPase}$ (□), phosphate transport, $q_{II:Pi\ trans}$ (Δ), and transport of metabolites $q_{II:metabolite\ trans}$ (○). The dashed line indicates the share of $q_{II:ATPase}$ used for maintenance and growth related ATP requirements.

**Transport fluxes**

The metabolic network analysis of the compartmented model also allows insight into the magnitude of the transmembrane fluxes, and thus into the activity of the transport protein. At $\mu = 0.1 \text{ hr}^{-1}$, the 93 to 435 mmol/C-mol/hr of the phosphate carrier and the associated ATP/ADP exchange carrier are by far the largest individual fluxes across the mitochondrial membrane (table IV). Costing one hydrogen for each phosphate translocation and one positive charge for each ATP/ADP exchange, thus a total of one proton, this amounts to an energy burden of about 0.3 to 1.3 mol ATP/C-mol biomass. These numbers exclude the flux required for the maintenance and growth related ATP requirements in the cytosol. With both $k_{ATP}$ and $m_{ATP}$ exclusively placed in the cytosol, an additional 760 and 590 mmol/C-mol/hr would have to be added to the transmembrane fluxes at $\mu = 0.1 \text{ hr}^{-1}$ in networks I and II respectively.
In comparison, proton dependent transporters involved in the anabolic processes account for a relative minor share of all energy consuming fluxes. Proton driven transport of metabolites totals between 12 and 31 mmol/C-mol/hr at \( \mu = 0.1 \ \text{hr}^{-1} \), the major contributor being pyruvate transport (table V). The two other significant metabolite fluxes are connected to the TCA cycle operation and the mitochondrial glutamate regeneration, namely the export of oxaloacetate and the transfer of valine and oxoisovalerate, the latter ones are summarized by doubling the valine import, to separate the oxoisovalerate export for anabolic purposes. The sum of all other metabolites account for a flux of 5-6 mmol/C-mol/hr. Summarizing, the total metabolite trafficking, excluding the antiports, poses an energy burden of about 36-93 mmol ATP per C-mol biomass formed, and will therefore have no measurable effect on the biomass yield. Taken alone, the trafficking of the anabolic precursors only accounts for a flux of about 40 to 50 mmol across the mitochondrial membrane per produced C-mol of biomass, independently of the growth rate, posing a potential burden of about 12-15 mmol ATP/C-mol of produced biomass.

**Fluxes of selected transport processes** (mmol/C-mol/hr)

<table>
<thead>
<tr>
<th>Network</th>
<th>( \Sigma q_{\text{precursor trans}} )</th>
<th>( q_{\text{PYR trans}} )</th>
<th>( q_{\text{OXACT trans}} )</th>
<th>( 2^* q_{\text{VAL trans}} )</th>
<th>( \Sigma q_{\text{metabolite trans}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethanol</td>
<td>I 4.6</td>
<td>3.9</td>
<td>0.3</td>
<td>3.0</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>II 4.6</td>
<td>3.9</td>
<td>0.3</td>
<td>3.0</td>
<td>12</td>
</tr>
<tr>
<td>acetate</td>
<td>I 4.9</td>
<td>4.1</td>
<td>0.3</td>
<td>3.2</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>II 4.6</td>
<td>3.9</td>
<td>0.3</td>
<td>3.0</td>
<td>12</td>
</tr>
<tr>
<td>glucose</td>
<td>I 4.4</td>
<td>23.3</td>
<td>0.2</td>
<td>2.8</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>II 4.4</td>
<td>22.9</td>
<td>0.2</td>
<td>2.8</td>
<td>30</td>
</tr>
</tbody>
</table>

*Table V*: calculated fluxes of selected transport processes for anabolic purposes at \( \mu = 0.1 \ \text{hr}^{-1} \), each causing an equal flux of protons into the mitochondria; \( \Sigma q_{\text{precursor trans}} \) denominates the sum of the transport fluxes of isoleucine, ornithine, oxoisovalerate, homocitrte, threonine, oxobutyrate, and ammonia.

Four antiporters were implemented in the metabolic networks I and II. Fluxes through the antiports total about 4 mmol/C-mol/hr for glucose grown cultures, where only the two oxodicarboxylate carriers (Odc1p or Odc2p) are expressed. However, ethanol or acetate as carbon source produce a flux of about 55 and 110 mmol/C-mol/hr respectively through the antiport systems with only slight differences between the networks I and II (table VI). With about equal share, the succinate-fumarate and the isocitrate-malate carrier contribute most to this high flux, mainly due to the cytosolic localization of parts of the glyoxylate cycle. If each antiport would be replaced by a pair of proton driven active transport systems, it would cause at \( \mu = 0.1 \ \text{hr}^{-1} \) a proton influx of about 110 and 220 mmol/hr per C-mol into the mitochondria for, respectively, ethanol and acetate grown biomass. This equals an energy burden of 0.33 and 0.66 mol ATP/C-mol. As a result, one would observe a drop in the biomass yield of respectively 5 and 8% for ethanol and acetate grown cells. Evidently, the use of antiports is highly beneficial under these conditions, a fact which is supported by the phenotypes of knock-out mutants for *SFC1* and *DIC1*, which are unable to grow on the C3-carbons ethanol and acetate (Fernandez et al., 1994; Palmieri et al., 1999b).
activity of antiport transport proteins (mmol/C-mol/hr)

<table>
<thead>
<tr>
<th>network</th>
<th>$q_{ODC1\ car}$</th>
<th>$q_{CTP1\ car}$</th>
<th>$q_{SFC1\ car}$</th>
<th>$q_{ODC2\ car}$</th>
<th>$2x\ \Sigma q_{\text{antiports}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethanol I</td>
<td>3.5</td>
<td>21.5</td>
<td>28.6</td>
<td>0.7</td>
<td>109</td>
</tr>
<tr>
<td>II</td>
<td>3.5</td>
<td>28.6</td>
<td>28.6</td>
<td>0.7</td>
<td>123</td>
</tr>
<tr>
<td>acetate I</td>
<td>24.6</td>
<td>51.1</td>
<td>28.9</td>
<td>0.7</td>
<td>211</td>
</tr>
<tr>
<td>II</td>
<td>25.2</td>
<td>57.2</td>
<td>28.6</td>
<td>0.7</td>
<td>223</td>
</tr>
<tr>
<td>glucose I</td>
<td>3.3</td>
<td>n/a</td>
<td>n/a</td>
<td>0.7</td>
<td>8</td>
</tr>
<tr>
<td>II</td>
<td>3.3</td>
<td>n/a</td>
<td>n/a</td>
<td>0.7</td>
<td>8</td>
</tr>
</tbody>
</table>

Table VI: calculated fluxes through the antiport carrier proteins oxoglutarate-malate ($ODC1$), (iso)citrate-malate ($CTP1$), succinate-fumarate ($SFC1$), and oxoglutarate-oxoadipate ($ODC2$) at $\mu = 0.1$ hr$^{-1}$; the last column showing the amount of proton flux required, if all antiports are replaced each by two individual active transport processes.

**Discussion and Conclusions**

The metabolic network analysis of a compartmented model allows quantification of the main metabolic transport processes. Under the conditions presented here, the main fluxes are the ones through the ATP/ADP antiports, the phosphate carrier, and the four implemented metabolite antiports. The analysis of the proton fluxes clearly indicates the energetic advantage of the use of this antiport system for the TCA cycle intermediates during growth on ethanol and acetate. Use of proton-motive-force for the shuttling of these components would add a significant additional energy burden. Yield calculations, using the derived energetic parameters without the antiports, results in a measurable drop of biomass formation in the magnitude of 5 - 8%.

The absence of a mitochondrial glutamate synthase or another ammonia binding reaction inside the mitochondria requires the use of a shuttle system to provide a mitochondrial nitrogen source. The most likely candidates for such a system are the aspartate/oxaloacetate and the valine/oxoisovalerate shuttle. The first shuttle would cause a significant outflow of oxaloacetate out of the mitochondria, which would exceed its inflow for anaplerotic purposes. However, only an oxaloacetate importer has been characterized for yeast (Palmieri et al., 1999a). We employed, therefore, the valine/oxoisovalerate shuttle, for which the appropriate transporters should exist according to the localization of the related pathways. Even with this system, we still found a small outward flux of oxaloacetate. This contradicts the established import function of the oxaloacetate carrier Oac1p, but might be solved with e.g. an additional antiport for the homocitrate, or a different biomass composition, which requires a higher amount of arginine relative to lysine.
The compartmented model clearly assigns different energetic values to cytosolic and mitochondrial ATP. The difference is depending on the efficiency \( \alpha \) of the \( F_0F_1 \)-ATPase, as can be seen from equation (8). For an ATPase working close to its optimum, the energetic cost of ten cytosolic ATP equals approximately that of thirteen mitochondrial ATP. On the other hand, the energetic value of NADH is independent of its localization in yeast, the only differences might be caused by the distinct concentrations in each compartment. This is distinctively different from other eukaryotic cell with a functional complex I. It allows in \( S. cerevisiae \) the unregulated exchange of NAD and NADH across the mitochondrial membrane via e.g. the acetaldehyde/ethanol shuttle (Bakker et al., 2000), which would result in the formation of an ATP-yield lowering substrate cycle in other cells. Therefore, a tight regulation of the alcohol dehydrogenases, as can be expected elsewhere, is not needed in baker’s yeast, possibly adding to its higher ethanol tolerance.

Another consequence of the energetic equivalence of cytosolic and mitochondrial NADH is found in the pentose phosphate pathway. The maximal ATP yield in \( S. cerevisiae \) from glucose is 27.5 mol ATP per mol glucose, when metabolised through the glycolysis. The yield of the PP-pathway is a mere 0.6 mol ATP/mol glucose lower, thus reaching 97% of the glycolysis’ yield, if GTP and ATP, as well as NADH and NADPH are considered equivalent. The difference for an organism with a proton-pumping complex I, is significantly larger, here, the PP-pathway yields only 93% of the ATP gained through the glycolysis. It is very well conceivable, that these differences are reflected in the difference of the redox states between the NAD/NADH and NADP/NADPH pairs.

In past publications (van Gulik and Heijnen, 1995; Vanrolleghem et al., 1996; Verduyn, 1992) dealing with the energetics of the yeast cell, it was generally understood that two complexes in the respiratory chain translocate six protons, and variably 3 or 4 protons were assumed to be required to generate one ATP at the \( F_0F_1 \)-ATPase. This was used to arrive at a theoretical P/O ratio of 2.0 or 1.5 respectively. However, taking the different energetic values of the ATP into account, and combining this with the latest scientific understanding of the respiratory chain and the ATPase, a new, lower theoretical maximal P/O value of 1.38 is derived based on cytosolic ATP.

The calculation of the energetic parameters indicate that both the respiratory chain as well as the \( F_0F_1 \)-ATPase are working very close to their predicted stoichiometry. The results indicate that energetic losses are more likely to be found in the respiratory chain. However, a separate determination of the individual efficiencies \( \alpha_1 \) and \( \alpha_2 \) of both processes is difficult due to the high sensitivity towards measuring errors in the original data. A more reliable determination of \( \alpha_2 \) could possibly be accomplished by the analysis of aerobic growth on glucose with variations on the nitrogen source, e.g. aspartate, which each will cause a very distinct amount for metabolic transport fluxes.

We were not able to select one of the two proposed metabolic networks as the definitive one based solely on flux analysis combined with the determination of the energetic parameters. Some indications seem to favour the network II over network I, such as the slightly better predictive capacity of the yield on acetate and the very high \( \alpha_1 \) in the first network. But these differences are not strong enough to exclude the validity of either network.
Energetic aspects of the compartmentation of the metabolism

The $k_{ATP}$ values decreases by 50 - 70% compared to previous un compartmented models due to the inclusion of the energy requirements for transport (Stückrath et al., 2001). This reduces the share of the non-network related energy metabolism to 40 and 20% for the C₅- and C₇-substrates respectively. More detailed metabolic networks, including further energy consuming reactions, e.g. trafficking of metabolites across other compartmental borders, transport of protein, and RNA processing, will continue to reduce the remaining growth related ATP-requirement $k_{ATP}$ for future models. However, as could be seen from the small value of the current transport terms, the contribution of further metabolic carriers will constitute only a minor share of the currently unspecified energy demand, lumped together in $k_{ATP}$ and $m_{ATP}$.

Acknowledgements

We would like to thank J. Pronk and W. van Gulik for critically reading the manuscript. The work was financed through the European framework IV program on cell factories and is part of the project ‘From gene to product in yeast: a quantitative approach’.

List of symbols

1. $D$\hspace{1cm}$D$ \hspace{0.5cm}$D$ \hspace{0.5cm}$D$ \hspace{0.5cm}$D$ \hspace{0.5cm}$D$ (hr⁻¹)
2. $k_{ATP}$\hspace{1cm}$k_{ATP}$ \hspace{0.5cm}$k_{ATP}$ \hspace{0.5cm}$k_{ATP}$ \hspace{0.5cm}$k_{ATP}$\hspace{0.5cm}$k_{ATP}$ growth related ATP requirements (mol/C-mol biomass)
3. $m_{ATP}$\hspace{1cm}$m_{ATP}$ \hspace{0.5cm}$m_{ATP}$ \hspace{0.5cm}$m_{ATP}$ \hspace{0.5cm}$m_{ATP}$ \hspace{0.5cm}$m_{ATP}$ cell maintenance expressed in ATP equivalents (mol/C-mol biomass/hr)
4. $q_{ATP-network}$\hspace{1cm}$q_{ATP-network}$ \hspace{0.5cm}$q_{ATP-network}$ specific rate (= flux) of ATP forming reaction $i$ (mol/C-mol biomass/hr)
5. $q_{X-trans}$\hspace{1cm}$q_{X-trans}$ \hspace{0.5cm}$q_{X-trans}$ \hspace{0.5cm}$q_{X-trans}$ flux through transport protein for compound $X$ (mol/C-mol biomass/hr)
6. $q_{H+\text{-trans}}$\hspace{1cm}$q_{H+\text{-trans}}$ \hspace{0.5cm}$q_{H+\text{-trans}}$ flux of active transporters excl. phosphate transport (mol/C-mol biomass/hr)
7. $q_{2e}$\hspace{1cm}$q_{2e}$ \hspace{0.5cm}$q_{2e}$ \hspace{0.5cm}$q_{2e}$ \hspace{0.5cm}$q_{2e}$ flux of electron pairs through the respiratory chain (mol/C-mol biomass/hr)
8. $\alpha_1$\hspace{1cm}$\alpha_1$ \hspace{0.5cm}$\alpha_1$ \hspace{0.5cm}$\alpha_1$ \hspace{0.5cm}$\alpha_1$ \hspace{0.5cm}$\alpha_1$ proton translocation efficiency of the respiratory chain
9. $\alpha_2$\hspace{1cm}$\alpha_2$ \hspace{0.5cm}$\alpha_2$ \hspace{0.5cm}$\alpha_2$ \hspace{0.5cm}$\alpha_2$ \hspace{0.5cm}$\alpha_2$ proton based ATP formation efficiency of the F₆F₁-ATPase
10. $\mu$\hspace{1cm}$\mu$ \hspace{0.5cm}$\mu$ \hspace{0.5cm}$\mu$ \hspace{0.5cm}$\mu$ \hspace{0.5cm}$\mu$ growth rate (hr⁻¹)

Sub-/superscripts

1. cyt\hspace{1cm}cyt \hspace{0.5cm}cyt \hspace{0.5cm}cyt \hspace{0.5cm}cyt \hspace{0.5cm}cyt \hspace{0.5cm}cyt relating to the cytosol
2. mit\hspace{1cm}mit \hspace{0.5cm}mit \hspace{0.5cm}mit \hspace{0.5cm}mit \hspace{0.5cm}mit \hspace{0.5cm}mit \hspace{0.5cm}mit relating to the mitochondria
Appendix: Metabolic network reactions

Amino acid synthesis

- cALA N-trans: PYR + GLM $\rightarrow$ OGL + ALA
- cARG syn: ORN + CARP + ATP + ASP $\rightarrow$ AMP + ARG + FUM + 3 H + Pi + PPi
- cASN syn: H2O + GLN + ATP + ASP $\rightarrow$ ADP + ASN + GLM + H + Pi
- cASP N-trans: OXACT + GLM $\rightarrow$ OGL + ASP
- cASP kin: 2 NADPH + 2 H + ATP + ASP $\rightarrow$ ADP + HSER + 2 NADP + Pi
- mIL-1 N-trans: OBU + PYR + NADPH + H + GLM $\rightarrow$ OGL + 1CO2 + H2O + ILE + NADP
- cLEU N-trans: NAD + H2O + GLM + OIV + ACCOA $\rightarrow$ OGL + CO2 + CoA + H + LEU + NADH
- cVAL N-trans: GLM + OIV $\rightarrow$ OGL + VAL
- mVAL N-trans: GLM + OIV $\rightarrow$ OGL + VAL
- cCARP syn: 2 H2O + GLN + CO2 + 2 ATP $\rightarrow$ 2 ADP + CARP + GLM + 3 H + Pi
- cCYS syn: H2S + CYC + HS + N4H + OBU
- cGLM N-lig: NH4 + GLM + ATP $\rightarrow$ ADP + GLN + H + Pi
- cGLM deh: NH4 + NADPH + H + OGL $\rightarrow$ GLM + H2O + NADP
- cGLY trans: THF + SER $\rightarrow$ GLY + H2O + METHF
- cHIS syn: PRPP + 2 NAD + 3 H2O + GLM + ATP $\rightarrow$ OGL + 6 H + 1HIS + 2 NADH + Pi + 2 Pi + AICAR
- cHCIT syn: H2O + OGL + ACCOA $\rightarrow$ CoA + H + HCIT
- cHCYS syn: H2S + HSER + ACCOA $\rightarrow$ ACT + CoA + 2 H + HCY
- cGLY syn: OAD + 2 NADPH + NAD + 2 GLM + ATP $\rightarrow$ 2 OGL + 1AMP + LYS + NADH + 2 NADP + Pi
- cMET syn: HCSY + MYTHF + H $\rightarrow$ MET + THF
- cORN syn: NADPH + H + 2 GLM + ATP $\rightarrow$ ADP + OGL + 1NADP + Pi + ORN
- mOAD syn: HCSY + NADP $\rightarrow$ CO2 + NADH + OAD
- mOIV syn: 2 PYR + NADPH + 2 H $\rightarrow$ OIV + CO2 + H2O + 1NADP
- cPHE syn: H + GLM + CHO $\rightarrow$ OGL + CO2 + H2O + PHE
- cPRO deh: 2 NADPH + 2 H + GLM + ATP $\rightarrow$ ADP + H2O + 2 NADP + Pi + PRO
- cSER syn: 3PG + NAD + H2O + GLM $\rightarrow$ OGL + H + NADH + Pi + SER
- cSHI path: 2 PEP + NADPH + EP + ATP $\rightarrow$ ADP + CHO + 1NADP + 4 Pi
- cTHR ald: THR $\rightarrow$ ACTAL + GLY
- mTHR deh: THR + H $\rightarrow$ NH4 + OBU
- cTHR syn: HSER + H2O + ATP $\rightarrow$ ADP + H + Pi + THR
- cTRP syn: SER + PRPP + GLN + CHO $\rightarrow$ CO2 + GAP + GLM + H + H2O + 2 Pi + PYR + TRP
- cTYR syn: NADP + GLM + CHO $\rightarrow$ OGL + CO2 + NADPH + 1TYR

Biomass formation

- bio-A1: 0.0176 metal + 0.00385 DNA + 0.13456 LIPID + 0.00188 SO4 + 0.05185 RNA + 0.29537 CARBYHD + 0.51428 PROT + 0.0163 Pi + 0.06329 H2O $\rightarrow$ 1biom-A1
- bio-A1 + k*: 0.0176 metal + 0.00385 DNA + 0.13456 LIPID + 0.00188 SO4 + 0.05185 RNA + 0.29537 CARBYHD + 0.51428 PROT + 0.0163 Pi + 0.06329 H2O + k*ATPM ATP + k*ATPM ADP + k*ATPM Pi + biom-A1
- biom-E1: 0.01799 metal + 0.00375 DNA + 0.15073 LIPID + 0.00186 SO4 + 0.04611 RNA + 0.32015 CARBYHD + 0.47917 PROT + 0.0196 Pi + 0.06574 H2O $\rightarrow$ biom-E1
- biom-E1 + k*: 0.01799 metal + 0.00375 DNA + 0.15073 LIPID + 0.00186 SO4 + 0.04611 RNA + 0.32015 CARBYHD + 0.47917 PROT + 0.0196 Pi + k*ATPM ATP + k*ATPM ADP + k*ATPM Pi + biom-E1
- biom-GA: 0.0178 metal + 0.00353 DNA + 0.10136 LIPID + 0.0013 SO4 + 0.05037 RNA + 0.38656 CARBYHD + 0.45874 PROT + 0.00218 Pi + 0.06049 H2O $\rightarrow$ biom-GA
- biom-GA + k*: 0.0178 metal + 0.00353 DNA + 0.10136 LIPID + 0.0013 SO4 + 0.05037 RNA + 0.38656 CARBYHD + 0.45874 PROT + 0.00218 Pi + k*ATPM ATP + k*ATPM ADP + k*ATPM Pi + biom-GA

C-1 metabolism

- cDHF red: DHF + NADPH + H $\rightarrow$ NADP + THF
- cMETHF deh: NADPH + H + FTHF $\rightarrow$ H2O + METHF + NADP
- cMETHF red: NADPH + METHF + H $\rightarrow$ MYTHF + NADP
catabolism
\[ \text{cSO4 ass} \quad \text{SO}_4^2- + 4 \text{NADPH}_c + 3 \text{H}_c + 2 \text{ATP}_c \rightarrow \text{ADP}_c + \text{AMP}_c + 2 \text{H}_2 \text{O}_c + 4 \text{NADP}_c + \text{Pi}_c + \text{PPI}_c + \text{H}_2\text{S}_c \]

diffusion
\[ \text{eCO2<->cCO2} \quad \text{CO}_2_c \rightarrow \text{CO}_2_e \]
\[ \text{eCO2<->mCO2} \quad \text{CO}_2_m \rightarrow \text{CO}_2_e \]
\[ \text{cO2->mO2} \quad \text{O}_2_c \rightarrow \text{O}_2_m \]
\[ \text{eH2O<->mH2O} \quad \text{H}_2\text{O}_m \rightarrow \text{H}_2\text{O}_e \]
\[ \text{eO2->cO2} \quad \text{O}_2_e \rightarrow \text{O}_2_c \]
\[ \text{eH2O<->cH2O} \quad \text{H}_2\text{O}_e \rightarrow \text{H}_2\text{O}_c \]

gluconeogenesis
\[ \text{cF16P phos} \quad \text{F16P}_c + \text{H}_2\text{O}_c \rightarrow \text{F6P}_c + \text{Pi}_c \]
\[ \text{cCIT ly a} \quad \text{ICIT}_c \rightarrow \text{GLYO}_c + \text{SUC}_c \]
\[ \text{cMAL syn} \quad \text{H}_2\text{O}_c + \text{GLYO}_c + \text{ACCoA}_c \rightarrow \text{CoA}_c + \text{H}_c + \text{MAL}_c \]
\[ \text{cPEP ck} \quad \text{OXACT}_c + \text{ATP}_c \rightarrow \text{ADP}_c + \text{CO}_2_c + \text{PEP}_c \]

glycolysis, lower
\[ \text{cEnol} \quad 2\text{PG}_c \rightarrow \text{H}_2\text{O}_c + \text{PEP}_c \]
\[ \text{cGAP deh} \quad \text{Pi}_c + \text{NAD}_c + \text{GAP}_c \rightarrow \text{H}_c + \text{NADH}_c + 13\text{PG}_c \]
\[ \text{c13PG kin} \quad 13\text{PG}_c + \text{ADP}_c \rightarrow \text{ATP}_c + 3\text{PG}_c \]
\[ \text{c3PG mut} \quad 3\text{PG}_c \rightarrow 2\text{PG}_c \]
\[ \text{cPYR kin} \quad \text{PEP}_c + \text{H}_c + \text{ADP}_c \rightarrow \text{ATP}_c + \text{PYR}_c \]

glycolysis, upper
\[ \text{cF16P ald} \quad \text{F16P}_c \rightarrow \text{GAP}_c + \text{DHAP}_c \]
\[ \text{cG6P iso} \quad \text{G6P}_c \rightarrow \text{F6P}_c \]
\[ \text{cHX kin} \quad \text{GLUC}_c + \text{ATP}_c \rightarrow \text{ADP}_c + \text{G6P}_c + \text{H}_c \]
\[ \text{cPF kin} \quad \text{F6P}_c + \text{ATP}_c \rightarrow \text{ADP}_c + \text{H}_c + \text{F16P}_c \]
\[ \text{cTP iso} \quad \text{DHAP}_c \rightarrow \text{GAP}_c \]

intracellular transport
\[ \text{ADP c<-> ATP m} \quad \text{ATP}_m + \text{ADP}_c \rightarrow \text{ADP}_m + \text{ATP}_c \]
\[ \text{ACTA lm car} \quad \text{ACTAL}_c \rightarrow \text{ACTAL}_m \]
\[ \text{ACTm imp} \quad \text{H}_c + \text{ACT}_c \rightarrow \text{ACT}_m + \text{H}_m \]
\[ \text{NH4m car} \quad \text{NH}_4_m + \text{H}_c \rightarrow \text{H}_m + \text{NH}_4_c \]
\[ \text{GPP c<-> ATP m} \quad \text{MAL}_m + \text{H}_c + \text{CIT}_c \rightarrow \text{CIT}_m + \text{H}_m + \text{MAL}_c \]
\[ \text{ETOH lm car} \quad \text{ETOH}_c \rightarrow \text{ETOH}_m \]
\[ \text{HCI2m car} \quad \text{HClIT}_c + 4 \text{H}_c \rightarrow 4 \text{H}_m + \text{HClIT}_m \]
\[ \text{CTP2m car} \quad \text{MAL}_m + \text{ICIT}_c + \text{H}_m \rightarrow \text{H}_m + \text{ICIT}_c + \text{MAL}_m \]
\[ \text{ILEm car} \quad \text{ILE}_m + \text{H}_c \rightarrow \text{H}_m + \text{ILE}_c \]
\[ \text{ORNm} \quad \text{ORN}_m + \text{H}_c \rightarrow \text{H}_m + \text{ORN}_c \]
\[ \text{OXACTm exp} \quad \text{OXACT}_m + \text{H}_m \rightarrow \text{H}_c + \text{OAXACT}_c \]
\[ \text{OXACTm imp} \quad \text{OXACT}_c + 3 \text{H}_c \rightarrow 3 \text{H}_m + \text{OAXACT}_m \]
\[ \text{OBUm car} \quad \text{OBU}_c + \text{H}_c \rightarrow \text{H}_m + \text{OBU}_c \]
\[ \text{OGLm exp} \quad \text{OGL}_m + \text{OGL}_c \rightarrow \text{OGL}_c + \text{H}_m \]
\[ \text{ODC1m car} \quad \text{MAL}_c + \text{OGL}_m \rightarrow \text{OGL}_c + \text{MAL}_m \]
\[ \text{ODC2m car} \quad \text{OAD}_m + \text{OGL}_m \rightarrow \text{OGL}_c + \text{OAD}_c \]
\[ \text{OIVm car} \quad \text{OIV}_m \rightarrow \text{OIV}_c \]
\[ \text{Pi m car} \quad \text{Pi}_c + 2 \text{H}_c \rightarrow 2 \text{H}_m + \text{Pi}_m \]
\[ \text{Hm slip} \quad \text{H}_c \rightarrow \text{H}_m \]
\[ \text{PYRm car} \quad \text{PYR}_c + 2 \text{H}_c \rightarrow 2 \text{H}_m + \text{PYR}_m \]
\[ \text{SFCm car} \quad \text{SUC}_c + \text{FUM}_m \rightarrow \text{FUM}_c + \text{SUC}_m \]
\[ \text{THRm car} \quad \text{THR}_c + \text{H}_c \rightarrow \text{H}_m + \text{THR}_m \]
\[ \text{VALm imp} \quad \text{VAL}_c + \text{H}_c \rightarrow \text{H}_m + \text{VAL}_m \]

lipid synthesis
\[ \text{cACCoA carb} \quad \text{H}_2\text{O}_c + \text{CO}_2_c + \text{ATP}_c + \text{ACCoA}_c \rightarrow \text{ADP}_c + 2 \text{H}_c + \text{Pi}_c + 1\text{MACoA}_c \]
\[ \text{cSAH hyd} \quad \text{SAH}_c + \text{H}_2\text{O}_c \rightarrow \text{H}_c + \text{HCYS}_c + \text{A}_c \]
\[ \text{cAVFA form} \quad 4.4 \text{PLLM-CoA}_c + 1.4 \text{PLM-CoA}_c + 1.7 \text{OLE-CoA}_c + \text{STE-CoA}_c \rightarrow 8.5 \text{avFA-CoA}_c \]
\[ \text{cAVPL form} \quad 4 \text{PHD-ETA}_c + 11 \text{PHD-CHO}_c + 3 \text{PHD-SER}_c \rightarrow 18 \text{avPL}_c \]
\[ \text{cFAT form} \quad \text{avFA-CoA}_c + \text{PHD}_c + \text{H}_2\text{O}_c \rightarrow \text{CoA}_c + \text{Pi}_c + \text{FAT}_c \]
cGOH3P trans 2 avFA-CoA, + GOH3P, -> 2 CoA, + PHD

cGOH3P deh DHAP, + NADH, + H, -> NAD, + GOH3P

mGOH3P deh DHAP, + NADH, + H, -> NAD, + GOH3P

cMET Atrans MET, + 2 H2O, + ATP, -> H2, + 3 Pi, + SAM

cPLM lig PLM-CoA, + H2O, + CoA, + ATP, -> AMP, + H2, + 2 Pi, + 1PLM-CoA

cPLM desat PLM-CoA, + O2, + NADPH, + H, -> 2 H2O, + NADP, + 1PLM-CoA

cPLM sym 7 MACoA, + 14 NADPH, + 20 H, + ACCoA, -> 7 CO2, + 8 CoA, + 6 H2O, + 14 NADP, + PLM

cPHD-C trans PHD, + H2O, + CTP, -> 2 Pi, + CMP-DGOH

cPHD-EA mtrans 3 SAM, + PHD-ETA, -> 3 H, + PHD-CHO, + 3 SAH

cPHD-SER dcarb PHD-SER, -> CO2, + PHD-ETA

ccPHD-SER syn CMP-DGOH, + SER, -> CMP, + PHD-SER

cSTE desat STE-CoA, + O2, + NADPH, + H, -> 2 H2O, + NADP, + 1OLE-CoA

cSTE ligase STE, + H2O, + CoA, + ATP, -> AMP, + H, + 2 Pi, + 1STE-CoA

cSTE sym 8 MACoA, + 16 NADPH, + 23 H, + ACCoA, -> 8 CO2, + 9 CoA, + 7 H2O, + 16 NADP, + STE

macromolecule synthesis
cAvAA form 0.2382 ORN, + 0.73257 VAL, + 0.19598 TYR, + 0.06492 TRP, + 0.55657 THR, + 0.5328 SER, + 0.42243 PRO, + 0.37583 PHE, + 0.11378 MET, + 0.65699 LYS, + 0.8008 LEU, + 0.58939 ILE, + 0.19258 HIS, + 0.88918 GLY, + 0.10214 GLM, + 0.977695 ALA, + 0.38557 ARG, + 0.40849 ASN, + 0.5199 ASP, + 01378 SYS, + 0.52618 GLN, -> 10 avAA

cCARBH4yd H2O, + G6P, + ATP, -> ADP, + H, + 2 Pi, + 6 CARBH4yd

cDNA poly 2 GTP, + 3 UTP, + NADPH, + 3 MTHF, + H, + 2 CTP, + 3 ATP, -> H2O, + NADP, + 9.8 DNA, + Ppi, + 3 DHF

cLipid form 0.45 avPL, + 0.55 FAT, -> 47.22 LIPID

cPROT poly 2 H2O, + 3 ATP, + avAA, -> 2 ADP, + AMP, + 4 H, + 2 Pi, + 4.810233 PROT, + PPI

cRNA syn 2 GTP, + 3 UTP, + 2 CTP, + 3 ATP, -> 9.5 RNA, + PPI

nucleotide synthesis
cA kin A, + ATP, -> ADP, + AMP, + H

cAMP kin ATP, + AMP, -> 2 ADP

cAMP synth ATP, + AMP, -> 2 ADP

cAMP syn IMP, + ATP, + ASP, -> ADP, + AMP, + FUM, + 2 H, + Pi

cCTP syn UTP, + H2O, + GLN, + ATP, -> ADP, + CTP, + GLM, + 2 H, + Pi

cCMP syn CTP, + ATP, -> ADP, + GDP

cGMP syn NAD, + IMP, + 2 H2O, + GLN, + ATP, -> AMP, + GLM, + GMP, + 4 H, + NADH, + PPi

cGMP kin GMP, + ATP, -> ADP, + GDP

cIMP syn AICAR, + FTHF, -> H2O, + IMP, + THF

cGDP kin GDP, + ATP, -> ADP, + GTP

cUDP kin UDP, + ATP, -> ADP, + UTP

cCDP kin CDP, + ATP, -> ADP, + CTP

cPRPP syn RIB5P, + ATP, -> AMP, + H, + PRPP

cAICAR syn PRPP, + 2 H2O, + GLY, + 2 GLN, + FTHF, + CO2, + 4 ATP, + ASP, -> 4 ADP, + FUM

cPiP syn PPI, + H2O, -> 2 Pi

cUMP syn PRPP, + 0.5 O2, + CARN, + ASP, -> CO2, + 2 H2O, + Pi, + UMP, + PPI

cUMP kin UMP, + ATP, -> ADP, + UDP

oxidative phosphorylation
mATPase Pi, + 10/3 H, + ADP, -> ATP, + 7/3 H + H2O

mFAD reg 0.5 O2, + FADH2, -> FAD, + H2O

mFADH2 deh 0.5 O2, + 5.66 H, + FADH2, -> FAD, + 5.66 H, + H2O

mNAD reg 0.5 O2, + NADH, + H, -> H2O, + NAD

mNADH deh 0.5 O2, + NADH, + H, -> 4.66 H, + H2O + NAD

mNAD deh 0.5 O2, + 6.66 H, -> 5.66 H, + H2O + NAD

pentose phosphate pathway
cG6P deh 2 NADP, + H2O, + G6P, -> CO2, + 2 H, + 2 NADPH, + RIB5P
cRIBU iso RIB5P, -> RIB5P
cRIBU ep RIB5P, -> XYL5P

cTA1 SED7P, + GAP, -> F4P, + F6P

120
Energetic aspects of the compartmentation of the metabolism

cTK1  XYL5P<sub>c</sub> + RIB5P<sub>c</sub> -> GAP<sub>c</sub> + SED7P<sub>c</sub>
cTK2  XYL5P<sub>c</sub> + E4P<sub>c</sub> -> F6P<sub>c</sub> + GAP<sub>c</sub>

pyruvate branchpoint

cACTAL deh (NAD)  NAD<sub>c</sub> + H2O<sub>c</sub> + ACTAL<sub>c</sub> -> ACT<sub>c</sub> + 2 H<sub>c</sub> + NADH<sub>c</sub>
mACTAL deh (NAD)  NAD<sub>m</sub> + H2O<sub>m</sub> + ACTAL<sub>m</sub> -> ACT<sub>m</sub> + 2 H<sub>m</sub> + NADH<sub>m</sub>
cACTAL deh (NADP)  NADP<sub>c</sub> + H2O<sub>c</sub> + ACTAL<sub>c</sub> -> ACT<sub>c</sub> + 2 H<sub>c</sub> + NADPH<sub>c</sub>
cACCoA syn  H2O<sub>c</sub> + CoA<sub>c</sub> + ATP<sub>c</sub> + ACT<sub>c</sub> -> ACCoA<sub>c</sub> + AMP<sub>c</sub> + H<sub>c</sub> + 2 Pi<sub>c</sub>
mACCoA syn  H2O<sub>m</sub> + CoA<sub>m</sub> + ATP<sub>m</sub> + ACT<sub>m</sub> -> ACCoA<sub>m</sub> + AMP<sub>m</sub> + H<sub>m</sub> + 2 Pi<sub>m</sub>
cETOH deh  NADH<sub>c</sub> + H<sub>c</sub> + ACTAL<sub>c</sub> -> ETOH<sub>c</sub> + NAD<sub>c</sub>
mETOH deh  NADH<sub>m</sub> + H<sub>m</sub> + ACTAL<sub>m</sub> -> ETOH<sub>m</sub> + NAD<sub>m</sub>
cLAC deh  NAD<sub>c</sub> + LAC<sub>c</sub> -> H<sub>c</sub> + NADH<sub>c</sub> + PYR<sub>c</sub>
cPYR-carb  PYR<sub>c</sub> + H2O<sub>c</sub> + CO2<sub>c</sub> + ATP<sub>c</sub> -> ADP<sub>c</sub> + 2 H<sub>c</sub> + Pi<sub>c</sub> + 10XACT<sub>c</sub>
cPDC  PYR<sub>c</sub> + H<sub>c</sub> -> ACTAL<sub>c</sub> + CO2<sub>c</sub>
mPYR deh  PYR<sub>m</sub> + NAD<sub>m</sub> + CoA<sub>m</sub> -> ACCoA<sub>m</sub> + CO2<sub>m</sub> + NADH<sub>m</sub>

TCA cycle

cACON 1  CIT<sub>c</sub> -> H2O<sub>c</sub> + AC0<sub>c</sub>
mACON 1  CIT<sub>m</sub> -> H2O<sub>m</sub> + AC0<sub>m</sub>
cACON 2  AC0<sub>c</sub> + H2O<sub>c</sub> -> ICIT<sub>c</sub>
mACON 2  AC0<sub>m</sub> + H2O<sub>m</sub> -> ICIT<sub>m</sub>
cCIT syn  OXACT<sub>c</sub> + H2O<sub>c</sub> + ACCoA<sub>c</sub> -> CIT<sub>c</sub> + CoA<sub>c</sub> + H<sub>c</sub>
mCIT syn  OXACT<sub>m</sub> + H2O<sub>m</sub> + ACCoA<sub>m</sub> -> CIT<sub>m</sub> + CoA<sub>m</sub> + H<sub>m</sub>
cFUM hy  H2O<sub>c</sub> + FUM<sub>c</sub> -> MAL<sub>c</sub>
mFUM hy  H2O<sub>m</sub> + FUM<sub>m</sub> -> MAL<sub>m</sub>
mCIT deh-NAD  NAD<sub>m</sub> + ICIT<sub>m</sub> -> OGL<sub>m</sub> + CO2<sub>m</sub> + NADH<sub>m</sub>
mCIT deh-NADP  NADP<sub>m</sub> + ICIT<sub>m</sub> -> OGL<sub>m</sub> + CO2<sub>m</sub> + NADPH<sub>m</sub>
cMAL deh  NAD<sub>c</sub> + MAL<sub>c</sub> -> H<sub>c</sub> + NADH<sub>c</sub> + OXACT<sub>c</sub>
mMAL deh  NAD<sub>m</sub> + MAL<sub>m</sub> -> H<sub>m</sub> + NADH<sub>m</sub> + OXACT<sub>m</sub>
mOGL deh  NAD<sub>m</sub> + CoA<sub>m</sub> + OGL<sub>m</sub> -> CO2<sub>m</sub> + NADH<sub>m</sub> + SUCCOA<sub>m</sub>
mSUC deh  SUC<sub>m</sub> + FAD<sub>m</sub> -> FADH2<sub>m</sub> + FUM<sub>m</sub>
mSUCCOA syn  SUCCOA<sub>m</sub> + Pi<sub>m</sub> + ADP<sub>m</sub> -> ATP<sub>m</sub> + CoA<sub>m</sub> + SUC<sub>m</sub>

transport

eACT trans  H<sub>c</sub> + ACT<sub>c</sub> -> ACT<sub>c</sub> + H<sub>c</sub>
eATPase  H2O<sub>c</sub> + ATP<sub>c</sub> -> ADP<sub>c</sub> + H<sub>c</sub> + Pi<sub>c</sub>
eETOH trans  ETOH<sub>c</sub> -> ETOH<sub>c</sub>
eGLUC diff  GLUC<sub>c</sub> -> GLUC<sub>c</sub>
eGOH trans  H<sub>c</sub> + GOH<sub>c</sub> -> GOH<sub>c</sub> + H<sub>c</sub>
eLAC trans  LAC<sub>c</sub> + H<sub>c</sub> -> H<sub>c</sub> + LAC<sub>c</sub>
eMET imp  metal<sub>c</sub> -> metal<sub>c</sub>
eNH4 trans  NH4<sub>c</sub> + H<sub>c</sub> -> H<sub>c</sub> + NH4<sub>c</sub>
ePi trans  Pi<sub>c</sub> + 2 H<sub>c</sub> -> 2 H<sub>c</sub> + Pi<sub>c</sub>
ePYR trans  PYR<sub>c</sub> + 2 H<sub>c</sub> -> 2 H<sub>c</sub> + PYR<sub>c</sub>
eSO4 trans  SO4<sub>c</sub> + 3 H<sub>c</sub> -> 3 H<sub>c</sub> + SO4<sub>c</sub>
References


Energetic aspects of the compartmentation of the metabolism


Chapter 6

Metabolic flux analysis of a triose-phosphate isomerase deletion mutant

submitted as

Metabolic flux analysis of a triose-phosphate isomerase deletion mutant of *S. cerevisiae*: Testing assumptions about metabolic networks by combining cofactor- and 13C-labelling balances

H.C. Lange, W. van Winden, D. Schipper, J.J. Heijnen
Metabolic flux analysis of a triose-phosphate isomerase deletion mutant

Abstract

Metabolic flux analysis (MFA) based on either the cofactor balances or $^{13}$C-labelling based isotopeomer balances were performed for the yeast *Saccharomyces cerevisiae* CEN.PK 113-7D and its isogenic triose-phosphate isomerase deletion (*tpiΔ*) mutant. Carbon-limited chemostat experiments at a dilution rate of $D = 0.1\text{hr}^{-1}$ provided the measured exchange fluxes against which a number of alternative metabolic networks were tested using the cofactor based MFA, allowing the elimination of several network options for the mutant based on statistical rejection. Flux calculations based on the $^{13}$C-spectra obtained from 2D [$^{13}$C, $^1$H] COSY-NMR spectrometry provided an alternative approach to get the metabolic fluxes. By comparing the results for both methods the number of valid networks could be reduced further. Furthermore, obtained network options were tested for their ability to describe both the wild type and the mutant data. This approach allowed to identify a network with a NADP coupled cytosolic acetaldehyde dehydrogenase, the pyruvate dehydrogenase, and the pyruvate decarboxylase as the most likely option out of 12 proposed networks for the wild type. Either one of two possible methylglyoxal pathways or a NADP dependent glycerol dehydrogenase were required to fit the NADP balance in the mutant. The activity of the methylglyoxal pathway in the *tpiΔ* mutant could neither be confirmed nor excluded with certainty and combination of both options resulted in the best fit between both MFA methods for the mutant data.
Introduction

Advances in microbiological tools allow the directed change of the genetic code of microorganisms. Here, a precise understanding of the metabolism is required to obtain the desired results. Metabolic flux analysis (MFA) has established itself as a valuable mathematical method to describe the activities in the cell’s reaction network (Christensen and Nielsen, 1999). It allows the quantification of the metabolic fluxes and thus can be used for hypothesis testing, prediction of metabolic changes, and as basis for a dynamic model of the cell’s metabolism, itself a prerequisite for an improved understanding of the metabolic network’s functioning.

Two different approaches are commonly used for the quantification of metabolic fluxes. The traditional metabolic flux analysis relies on the use of metabolite and co-factor balances. Its limitations for the determination of parallel and bidirectional pathways can be overcome by the $^{13}$C-labelling based MFA, which utilises the labelling patterns of metabolites to determine the distribution of fluxes through the network. The latter method requires the use of nuclear-magnetic resonance spectrometry (NMR) or mass spectrometry and is more labour intensive and not completely without uncertainties (van Winden et al., 2001b).

The yeast *Saccharomyces cerevisiae* has been the subject of numerous studies. Its common use for industrial applications, such as ethanol production, dough raising, heterologous expression of proteins, the relative ease of its cultivation, and the availability of the complete genome sequence have made it a popular organism for quantitative studies of the cell’s metabolism. Previous studies included the evaluation of the metabolic fluxes of yeast growing on various substrates (de Jong-Gubbels et al., 1995), using a partly compartmented metabolic network (Nissen et al., 1997), the determination of the energetic parameters during aerobic growth (Lange and Heijnen, 2002), and the prediction of the behaviour of deletion mutants (Stückrath et al., 2001).

$^{13}$C-labeling based MFA was recently applied to *S. cerevisiae* CEN.PK117-7D by Gombert et al. (Gombert et al., 2001) and by Maheimo et al. (Maheimo et al., 2001). Both used a model of the central carbon metabolism consisting of separate cytosolic and mitochondrial compartments. Gombert et al. investigated the fluxes in the wild type strain growing on either high or low glucose concentrations using aerobic batch and chemostat fermentations, respectively. They applied GC-MS to measure the $^{13}$C-labelling patterns, with which the flux distribution was calculated. Maheimo et al. applied $^{13}$C,$^1$H-NMR measurements to compare the fluxes in the wild type strain growing aerobically and anaerobically in shake flask cultures, thus obtaining valuable information on the essentially anaerobic metabolism.

In the presented work, the same wild type yeast will be compared with it’s isogenic triosephosphate isomerase deletion (tpiΔ) mutant, both cultivated in aerobic chemostat fermentations. The tpiΔ mutant, which cannot interconvert the glycolytic intermediates dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate (GAP), has
attracted the interest of researchers due to its capacity to produce high levels of glycerol (Compagno et al., 1996; Compagno et al., 2001). However, glycerol yields were lower than initially anticipated, leading to speculation on the fate of the produced DHAP. A potential solution would be an increased flux through the methylglyoxal pathway, which contributes less than 1% to the glucose catabolism in wild type yeast (Martins et al., 2001).

Two ways of MFA, a cofactor-based one and one using the $^{13}$C-labelling patterns of the amino acids obtained by $^{13}$C-NMR, were applied to determine the flux distribution of the wild type yeast and its isogenic $tpi1\Delta$ mutant in a compartmented metabolic network. The strengths and weaknesses of the different modelling approaches will be discussed. The focus of the presented study will be on the glycolysis, the pentose phosphate pathway, and the TCA cycle; the main pathways for the most common substrate, glucose and similar sugars. Here, many questions, e.g. the split ratio between pentose phosphate pathway and glycolysis, remain to be answered despite a wealth of information and models for these central pathways. Combining the results of both MFA methods, the different options for possible metabolic networks are evaluated. The aim of this work is therefore not only to yield the best estimation of the fluxes for a given metabolic reaction network, but also to identify the most likely network itself.

**Material and Methods**

**Organism**

The haploid, prototrophic *Saccharomyces cerevisiae* strain CEN.PK-113.7D was used during this study. Its isogenic triose-phosphate isomerase deletion ($tpi1\Delta$) mutant was obtained as described by Compagno et al. (Compagno et al., 2001).

**Culturing conditions**

Aerobic, carbon-limited chemostat cultivations were performed in 2-litre Applicon fermentors (Applicon Dependable Instruments, Schiedam, The Netherlands) at pH 5.0, 30°C and a stirrer speed of 800 rpm. Cultivation proceeded at a constant dilution rate of 0.1 hr$^{-1}$ for the wild type and 0.082 hr$^{-1}$ for the mutant on a defined mineral medium according to Verduyn et al. with 3 g/l glucose and 0.3 g/l ethanol as dual carbon source; ammonia provided the required nitrogen (Verduyn et al., 1992). Inoculation was performed with 50 ml preculture grown overnight in 11 shake flasks on identical medium. The cultures were allowed to reach steady state, usually within 4-5 dilution times, before switching to medium containing 10% uniformly $^{13}$C-labelled glucose (D-glucose-U-$^{13}$C$_6$, CAS: 492-62-6, Cambridge Isotope Laboratories Inc., Andover, MA, USA).

Steady state conditions were routinely checked by measurements of the biomass concentrations through filtration on nitro-cellulose filters (pore size 0.45 µm, Gelman Science, Ann Arbor, MI, USA). Carbon balances closed within 5%.

Humidity in the off-gas was removed through a condenser (<2°C) and a Perma Pure dryer (type MD-110-48P, Perma Pure Inc, Toms River, N.J., USA). Subsequently, the O$_2$ and
CO₂ concentration were measured using a Servomex 1100A Oxygen Analyser (Taylor Servomex, Crowborough, UK) and a Beckman 864 infrared detector (Rosemount Analytics, Santa Clare, CA., USA), respectively. Gas flow rates were measured using a Saga Digital Flow Meter (Ion Science, Cambridge, England).

Biomass was harvested 35.4 hours after switching to labelled substrate, i.e. after replacing 97% of unlabelled biomass by ¹³C-labelled one. For the wild type strain, an additional, smaller sample was taken after 11.8 hours. The samples of fresh fermentation broth were directly centrifuged (6 min., 4800 rpm), the cell pellet was washed with 0.9% NaCl-solution and subsequently with demineralized water, before being frozen at -80°C. Prior to NMR analysis the biomass was lyophilized for 48 hours.

¹³C-NMR

The dried biomass was hydrolysed in 10 ml 6 N HCl for 16 hours at 110°C. After filtration and evaporation to dryness, the residue was dissolved in 10 ml 0.1 N HCl and the amino acids were adsorbed to an ion exchange resin (Dowex AG 50W X4, Dow Chemical, MI, USA) and washed with 0.1 N HCl. The amino acids were eluted with 4 N HCl. After evaporation the residue was dissolved in D₂O. Water extracts of cell components other than proteinogenic amino acids were prepared by heating the biomass in 5 ml H₂O to 90°C for 10 minutes with shaking. After centrifugation, the supernatant was lyophilized and dissolved in D₂O.

NMR measurements were performed at 600 MHz at 37 °C on a Bruker Avance 600 spectrometer (Bruker, Rheinstetten, Germany). The [¹³C,¹H] COSY experiment was the HSQC sequence by Bax and Pochapsky (Bax and Pochapsky, 1992) with gradients for artefact suppression. Folding in F1 was used for reducing the sweepwidth. The ¹³C carrier was set to 57.5 ppm and 2400 increments were recorded with an effective sweep width of 20 ppm (t₁max =398 ms). For the aromatic carbons the offset was 129.6 ppm and 512 increments were recorded with a sweepwidth in F1 of 3 ppm (t₁max =652 ms) The window function used before Fourier transformation were a cosine-bell shifted by p/3 in F2 and a sine-bell shifted by p/2.6. Two-dimensional sections along the ¹³C-dimension were made from the three-dimensional [¹³C,¹H] COSY NMR spectra. The resulting ¹³C-spectra were fitted and measurement errors were determined from the NMR-noise according to Van Winden et al. (van Winden et al., 2001a).

Metabolic flux calculations

Two modelling approaches were used to calculate the metabolic fluxes: flux analysis was performed using metabolite balances combined with either co-factor or ¹³C-isotopomer balances.

The first approach relied on the material balances of the reaction network including the cofactor balances for NAD(H) and NADP(H) which permitted the determination of certain parallel pathways within the cell. The use of a compartmented model allowed the distinction between mitochondrial and cytosolic NAD(P)H generation, thus imposing a further restriction on the metabolic network. Balancing and metabolic flux analysis was performed using the software MNA V2.0 (SpadIT, Nijmegen, The Netherlands). The
biomass composition of the wild type and the mutant were assumed to equal the average biomass composition on glucose given by Lange and Heijnen (Lange and Heijnen, 2001). Two different networks were employed, which differed with respect to the reactions related to the pyruvate bypass (Figure 1). In network I, acetaldehyde is transformed to citrate via acetaldehyde dehydrogenase, acetyl-CoA synthase, and citrate synthase in the cytosol, while in network II this pathway is mitochondrial using an NAD coupled acetaldehyde dehydrogenase (Lange and Heijnen, 2002). Thus the carbon skeleton enters the mitochondria in the form of citrate in network I and as acetaldehyde in network II. In both networks, the cytosolic sources of NADPH were the pentose phosphate pathway and potentially the cytosolic acetaldehyde dehydrogenase, for which isoenzymes with cofactor specificity for both, NAD and NADP, have been reported for yeast. Generally, the NADP dependent acetaldehyde dehydrogenase is considered more likely to be present than the NAD dependent isoenzyme. However, due to the significance of this assumption on the calculated fluxes in the tpiΔ mutant, modelling was performed for both cofactors in both networks. A further restriction was introduced by the use of the energy balance, which was based on the energetic parameter described in earlier work (Lange and Heijnen, 2002) (Table I). The same networks were used for the wild type and the tpiΔ mutant, with an additional proton-driven export for the glycerol implemented for the latter case.

**Figure 1**: central metabolism as modelled in the networks I with a solely cytosolic acetyl-CoA synthase and in network II with an additional acetaldehyde dehydrogenase and acetyl-CoA synthase in the mitochondrion.

The 13C-labelling based MFA was performed on the basis of the same compartmented model that was used for the metabolite balance based MFA, except for the cofactor balances for ATP, NADH and NADPH that were left out. An additional reaction in this network was the reaction from malate to pyruvate catalysed by the malic enzyme, which can not be determined through cofactor-based MFA. The reversible reactions of the glucose-6-phosphate (G6P) isomerase, the reactions of the non-oxidative branch of the
pentose phosphate pathway and mitochondrial transporters for pyruvate and oxaloacetate were modelled each as two separate reactions catalysing the forward and backward steps, respectively. A further difference in the model used for the $^{13}$C-labelling based MFA was the introduction of additional transketolase and transaldolase reactions that do not influence the energy and reduction status of the cell but that do influence the labelling pattern (van Winden et al., 2001b). Finally, a reaction was added leading from succinate back to itself. This stoichiometrically trivial reaction is introduced to account for the scrambling of the $^{13}$C-labelling between the top and bottom half of the succinate and fumarate molecules in the TCA-cycle which occurs due to their symmetry. A total of 15 additional reaction were thus added to the network.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Network I</th>
<th>Network II</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P/O_{cyt}$ (mol cyt. ATP/mol 2e-)</td>
<td>1.30</td>
<td>1.18</td>
</tr>
<tr>
<td>$k_{ATP}$ (mmol/Cmol)</td>
<td>300</td>
<td>170</td>
</tr>
<tr>
<td>$m_{ATP}$ (mmol/Cmol/hr)</td>
<td>49</td>
<td>43</td>
</tr>
</tbody>
</table>

Table 1: energetic parameters used for modelling purposes, parameter sets were derived from glucose and ethanol grown chemostat cultures of S. cerevisiae CEN.PK-113.7D and reported earlier (Lange and Heijnen, 2002)

The number of isotopomers of all the metabolites in a metabolic network is very large ($2^n$ for a single n-carbon compound). This leads to a large set of labelling balances needed to simulate the $^{13}$C-labelling distribution in the metabolites. This set can, however, be considerably reduced without losing any accuracy by simplifying the metabolic network prior to setting up the balances. This is done by omitting all metabolite pools that have only a single influx and lumping the fluxes leading to and from those so-called ‘linear and diverging nodes’ (van Winden et al., 2001). This lumping results in all biosynthetic fluxes being lumped into single outflows of the metabolites of the central carbon metabolism (figure 3). The only biomass component that has a double influx in the model is serine, which may be either directly derived from 3-phosphoglycerate or may be formed from glycine via the reversible serine hydroxymethyl transferase reaction.

The two networks used for simulating the $^{13}$C-labelling distribution in the wild type and in the $tpiA$ mutant differed in two respects. Firstly, the lumped reaction leading from fructose-6-phosphate (F6P) to two molecules of glyceraldehyde-3-phosphate (GAP) in the model of the wild type was replaced by a lumped reaction leading from fructose-6-phosphate to one molecule of GAP and one molecule of pyruvate for the mutant case (i.e. the integrated hexose cleavage and methylglyoxal bypass for the dihydroxyacetone phosphate (DHAP) moiety). Secondly, in contrast to the wild type, the mutant was assumed to have a lumped reaction leading from F6P to one molecule of GAP and one molecule of glycerol (i.e. the integrated hexose cleavage, reduction to glycerol, and excretion of the DHAP moiety).
For simulating the $^{13}$C-labelling pattern, the net fluxes entering the metabolic network as carbon substrates and leaving the network towards anabolism were obtained by balancing the measured net conversion rates using the non-lumped, detailed metabolic network from which the cofactor balances were omitted. These balanced rates were combined with the mass balances of the converging nodes that remained in the reduced metabolic model to constrain the fluxes in the $^{13}$C-labelling simulation. In the model of the wild type strain the total number of 44 fluxes were thus restricted by 26 constraints, and in the model of the tpiΔ mutant 45 fluxes were restricted by 27 constraints. In both cases, 18 parameters that fixed the remaining degrees of freedom had to be determined by fitting the simulated $^{13}$C-labelling data to the measured ones.

The simulation of the $^{13}$C-labelling distribution was implemented in Matlab (The MathWorks, Inc.) based on the approach proposed by Wicheert et al. (Wicheert et al., 1999). The obtained model was used for a linearly constrained nonlinear minimization of the residual of the simulated and the measured 2D [$^{13}$C, $^1$H] COSY NMR data in order to find the metabolic fluxes that yielded the minimal covariance weighted sum of squared residuals.

**Results**

**Wild type**

Both strains were grown in a chemostat culture on the same mixture of 10:1 (w/w) glucose and ethanol to allow comparison between the wild type *S. cerevisiae* CEN.PK-113.7D and its ethanol requiring isogenic tpiΔ mutant. The wild type reached a steady state biomass concentration of 1.70 g/l, resulting in a yield of 16.3 g biomass per C-mol substrate consumed. No byproducts were found in the supernatant of the wild type cultivation. Yields and biomass levels followed the expectations based on earlier experiments with pure substrates (Stückrath et al., 2001).

<table>
<thead>
<tr>
<th>specific exchange rates balanced with black box model</th>
<th>wild type</th>
<th>ΔTPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>glucose</td>
<td>-24.8</td>
<td>-32.4</td>
</tr>
<tr>
<td>ethanol</td>
<td>-9.7</td>
<td>-14.2</td>
</tr>
<tr>
<td>oxygen</td>
<td>-73.2</td>
<td>-84.0</td>
</tr>
<tr>
<td>carbon dioxide</td>
<td>68.0</td>
<td>83.1</td>
</tr>
<tr>
<td>glycerol</td>
<td>-</td>
<td>19.2</td>
</tr>
<tr>
<td>biomass</td>
<td>100.0</td>
<td>82.0</td>
</tr>
</tbody>
</table>

Table II: reconciled values of the main exchange fluxes for aerobic, carbon-limited chemostat cultivation of *S. cerevisiae*, grown at $D = 0.01$ h$^{-1}$ (wild type) and $D = 0.08$ h$^{-1}$ (tpiΔ mutant) based on elemental and electron balances.
Initial black box balancing of the exchange rates confirmed the validity of these measurements. The balanced rates are given in table II. Comparison between model predictions and measurements using the compartmented metabolic network with the cofactor balance and the energetic parameter from earlier work (table I) resulted in a good match for the wild type. Both networks were tried with a different cofactor specificity for the acetaldehyde dehydrogenase. Deviating from the maximal yield concept, the pyruvate bypass via the pyruvate decarboxylase was included in the metabolic network as option in addition to the pyruvate dehydrogenase. Hence, a total of twelve different networks were tested. Statistical evaluation and reconciliation of the measured data using cofactor and energy balances resulted in a maximal $X^2$ test value of 12.4 for a seven times overdetermined system, well below the 95% cutoff value of 14.1, thus a faulty system definition could be excluded in all cases (table III). However, variation 3 exhibited a negative flux through the pyruvate dehydrogenase and must therefore be excluded as possible option.

| Variation | Network I (I or II) | Acetaldehyde deh. | NADPH deh. | Pyruvate decarb. | df | h | pass | Biom | Gluc | Eth | G6P iso | G6P deh | PYR deh | PYR dec | PYR carb | PYR kin |
|-----------|---------------------|-------------------|------------|------------------|----|---|------|------|------|------|--------|---------|---------|---------|---------|----------|--------|
| 1         | x                   | x                 | x          |                  | 7  | 9.88 | y    | 4.19 | 1.00 | 0.352 | 0.379   | 0.352   | 0.721   | 0       | 0.224    | 1.13    |
| 2         | x                   | x                 | x          | x                | 6  | 4.24 | y    | 4.03 | 1.00 | 0.391 | 0.645   | 0.093   | 0.365   | 0.480   | 0.216    | 1.24    |
| 3         | x                   | x                 | x          | x                | 7  | 10.1 | y    | 3.89 | 1.00 | 0.393 | 0.924   | -0.173  | 0       | 0.971   | 0.208    | 1.35    |
| 4         | x                   | x                 | x          |                  | 7  | 8.41 | y    | 4.16 | 1.00 | 0.391 | 0.186   | 0.547   | 0.863   | 0       | 0.223    | 1.06    |
| 5         | x                   | x                 | x          | x                | 6  | 4.24 | y    | 4.03 | 1.00 | 0.391 | 0.209   | 0.530   | 0.345   | 0.353   | 0.216    | 1.09    |
| 6         | x                   | x                 | x          | x                | 7  | 9.21 | y    | 3.91 | 1.00 | 0.393 | 0.235   | 0.513   | 0       | 0.743   | 0.209    | 1.12    |
| 7         | x                   | x                 | x          |                  | 7  | 12.4 | y    | 4.03 | 1.00 | 0.392 | 0.301   | 0.428   | 0.886   | 0       | 0.22     | 1.09    |
| 8         | x                   | x                 | x          | x                | 6  | 4.24 | y    | 4.03 | 1.00 | 0.391 | 0.329   | 0.410   | 0.128   | 0.610   | 0.216    | 1.13    |
| 9         | x                   | x                 | x          | x                | 7  | 4.67 | y    | 3.99 | 1.00 | 0.391 | 0.335   | 0.406   | 0       | 0.750   | 0.213    | 1.14    |
| 10        | x                   | x                 | x          |                  | 7  | 11.1 | y    | 4.20 | 1.00 | 0.392 | 0.178   | 0.552   | 0.851   | 0       | 0.224    | 1.06    |
| 11        | x                   | x                 | x          | x                | 6  | 4.24 | y    | 4.03 | 1.00 | 0.391 | 0.209   | 0.530   | 0.136   | 0.563   | 0.216    | 1.09    |
| 12        | x                   | x                 | x          | x                | 7  | 4.72 | y    | 3.99 | 1.00 | 0.391 | 0.217   | 0.524   | 0       | 0.709   | 0.213    | 1.10    |

Table III: twelve tested metabolic network options for the wild type varying in the chosen network (I or II, inclusive energetic parameters), the cytosolic acetaldehyde dehydrogenase (NAD or NADPH dependent), and the use of the pyruvate dehydrogenase and the pyruvate decarboxylase. Furthermore, the degree of freedom (df) together with the $X^2$-test values (h) indicate the statistical acceptance (pass) of the measurement. The cutoff values for the 95% confidence interval are 14.1 and 12.6 for a degree of freedom of 7 and 6, respectively. Metabolic fluxes in moles per mol glucose are given for the main substrates, biomass, through glycolysis and pentose phosphate pathways, and around the pyruvate branch point.

The predicted split ratio between the flux through the G6P dehydrogenase to the one through the G6P isomerase varied from 0.09:0.65 to 0.55:018 depending on the choice of network. As expected, networks employing an NAD dependent acetaldehyde dehydrogenase have a higher split ratio. For the network I inclusion of both the pyruvate dehydrogenase and the pyruvate decarboxylase resulted in all cases in a much better statistical fit than with inclusion of only one of the two. In network II however, the difference between the networks with both reactions and networks with only the decarboxylase were minimal with a slight preference for the latter.
For the wild type *S. cerevisiae*, biomass harvested after 11.8 and 35.4 hours after switching to $^{13}$C-labelled substrate was analysed for 98 spectra of various metabolic endproducts, i.e. amino acids, trehalose and levulinic acid (a degradation product of glucose). Both sets of data were corrected to a label distribution at infinite time (van Winden et al., 2001a), resulting in a good match between the measuring sets of both biomass harvesting times. A notable exception was the $\alpha$-carbon atom of glycine which gave a clearly different spectrum at early and late harvesting times for unknown reasons.

The available duplicate measurements of spectra revealed some inconsistencies for the $\alpha$-carbon atom of methionine and the $\delta$-carbon atom of histidine. As these inconsistencies are considerably higher than the measurement error that was estimated from the spectra by the peakfitting software, other as yet unknown sources of measurement errors must be present. A tentative source of such errors is the dependence of relative intensities on the exact $^1$H-frequency where the two-dimensional $^{13}$C-section of the three-dimensional [$^{13}$C,$^1$H] spectrum is made (van Winden et al., 2001a).

Initial examination of the relative intensities (i.e. the relative areas of the various peaks) of the spectra confirmed the biosynthetic origins and pathways that were published by Maahéimo et al. (Maahéimo et al., 2001). By consequence, reasonably close correspondence could be found between the spectra of carbon atoms of the group alanine, valine and parts of isoleucine and leucine, the group aspartate, methionine and threonine and parts of isoleucine, the group glutamate, proline, arginine and a part of lysine, and the group of phenylalanine and tyrosine. Serine was found not to correspond to the part of tyrosine and phenylalanine that also stems from the trioses in the lower glycolysis. This suggests that serine is reversibly cleaved into glycine, as was assumed in the network model.

For the wild type the measured labelling pattern in the part of serine that ends up in glycine and the labelling pattern in glycine itself do not correspond well in the early harvested biomass. This could point at a possible additional route of glycine production from threonine by threonine aldolase. However, as was stated above, spectra of glycine in early and late harvested biomass of the wild type gave surprisingly different results.

The biosynthetic origin of alanine is still ambiguous. In order to decide which biosynthetic origin to choose in our model, we compared the spectra of alanine with those of (parts of) valine, leucine and isoleucine that all stem from mitochondrial pyruvate and to those of (parts of) tyrosine and phenylalanine that stem from phosphoenolpyruvate that is a precursor for cytosolic pyruvate. Although the differences were not very large, the spectra of alanine were most similar to the mitochondrial amino acids.

As was discussed by van Winden et al. (van Winden et al., 2001a) the spectra of some amino acids showed more peaks than may be expected from the splitting by directly neighbouring $^{13}$C-atoms alone. This is for example the case for the $\beta$- and $\delta$-carbon atoms of histidine and the 3rd carbon atom of levulinic acid. Separately fitting these additional peaks may lead to better discrimination of the fluxes. Therefore, we included these more detailed spectra in our measurement set.

For the wild type the measurement data set that was fitted by the $^{13}$C-labelling model
consisted of 322 relative intensities in 98 spectra of various carbon atoms. Due to the normalization of the areas for each spectrum, this means that 322-98=224 independent data points were available to fit the 18 degrees of freedom in the set of flux constraints. The metabolic flux calculation resulted in a covariance-weighted residual error between the measured and simulated spectra of 18996. This minimized squared error is much larger than the statistically expected value of 206 (224 independent data points minus 18 parameters). This may point at the presence of biochemical errors or omissions in the metabolic model and/or underestimation of the real measurement errors by assuming that the NMR-noise is the sole source of errors in our data. Figure 2 shows the parity plot of the measured and simulated $^{13}$C-labelling data. The largest outlier in this plot that has a measured value of 0.45 and a simulated value of 0.22 is the relative intensity of a set of peaks in the NMR spectrum of the β-carbon of histidine that largely overlap with another set of peaks. By consequence, the relative intensities of these peaks cannot very well be estimated separately. Beside this major and some minor outliers, the measured and simulated data generally agree quite well (figure 2). The large covariance-weighted residual error that is found is therefore rather the result of small values of the estimated measurement errors than of large absolute differences between the simulated and measured data.

![Parity plot](image)

**Figure 2:** Parity plot of the wild type data for *S. cerevisiae* CEN.PK113-7D with the simulated relative intensities plotted over the measured relative intensities.

The metabolic flux distribution based on the $^{13}$C-labelling allowed the quantification of parallel pathways and reversible fluxes without using the co-factor balances. This resulted in a flux pattern of the central metabolism, including the glycolysis, the PP-pathway, and the TCA cycle, derived independently from the one obtained through cofactor based MFA. Here, the split ratio between G6P dehydrogenase and its isomerase was determined as 0.32:0.42 (figure 3). The contribution of the pyruvate dehydrogenase was calculated as 0.50, leaving a significant proportion (0.15) of pyruvate to be metabolised via the pyruvate decarboxylase and thus forming acetyl-CoA via the pyruvate bypass. The flux
through the malate dehydrogenase (malic enzyme) was determined to be zero by the $^{13}$C-based MFA. The reversible fluxes in the G6P isomerase were calculated as 1.5 mol/mol glucose additionally to the net-flux of 0.39 towards fructose-6-phosphate (F6P). The exchange flux of C$_4$-components across the mitochondria were estimated to 1.3 mol in addition to the net-influx of 0.15.

![Metabolic flux analysis of a triose-phosphate isomerase deletion mutant](image)

**Figure 3:** flux pattern as determined by the $^{13}$C-labelling based MFA for the wild type *S. cerevisiae* CEN.PK113-7D grown at $D=0.1/hr-1$ in carbon limited chemostat culture, bidirectional fluxes in the metabolic network are indicated by double-sided arrows of which the closed head points in the direction of the net flux.

The confidence intervals of these fluxes in figure 3 cannot be calculated from the estimated measurement errors by applying the error propagation theory. The reason for this is that the estimated measurement errors cannot explain the observed residuals between the measured and simulated $^{13}$C-data, as was mentioned before, and the optimal fit therefore has to be rejected on statistical grounds. Still, not knowing the additional source of errors, the reliability of the presented flux values can be commented upon.
For this purpose, we calculated the matrix containing the sensitivities of the eighteen optimally fitted free parameters to the measured $^{13}$C-data. When analysing the singular values of this sensitivity matrix it was found that six of the eighteen were considerably (one >3700-fold and the rest >126000-fold) smaller than the other twelve. This means that in fact the optimal parameter values form a six-dimensional subspace in the parameter space, rather than a unique solution. This parameter-subspace was translated to a flux-subspace of the same dimension. Table IV shows the intervals within which the fluxes in figure 4 may be varied without significantly affecting the fit between the simulated and measured data set and without violating the constraint that all fluxes should have positive values.

<table>
<thead>
<tr>
<th>reaction</th>
<th>flux</th>
<th>minimum</th>
<th>maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s7p + g3p &gt; 2 p5p$</td>
<td>0</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>$f6p + g3p &gt; p5p + e4p$</td>
<td>0.25</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>$s7p + e4p &gt; f6p + p5p$</td>
<td>0</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>$f6p + p5p &gt; s7p + e4p$</td>
<td>0.11</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>$f6p + e4p &gt; e4p + f6p$</td>
<td>0</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>$s7p + p5p &gt; p5p + s7p$</td>
<td>0</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>$s7p + e4p &gt; e4p + s7p$</td>
<td>0</td>
<td>246.41</td>
<td></td>
</tr>
<tr>
<td>pyr (cyt) &gt; pyr (mit)</td>
<td>0.67</td>
<td>20.24</td>
<td></td>
</tr>
<tr>
<td>pyr (mit) &gt; pyr (cyt)</td>
<td>0</td>
<td>19.56</td>
<td></td>
</tr>
<tr>
<td>succ &gt; succ</td>
<td>0</td>
<td>&gt;1000</td>
<td></td>
</tr>
</tbody>
</table>

**Table IV:** Error margins for the most insensitive metabolic fluxes determined with $^{13}$C-labelling in the wild type strain, showing those fluxes that have absolute values larger than 0.005 and that can deviate more than 5% from their optimally estimated values.

From the table it is clear that the fluxes in the non-oxidative branch of the pentose phosphate pathway, the pyruvate exchange over the mitochondrial membrane and the succinate scrambling rate are not well defined. Their values can be freely varied without affecting the fit. As said, their true confidence intervals can only be determined when the true measurement errors are known.

Comparison of the results from the co-factor based MFA with the ones based on $^{13}$C-labelling showed the best fit for the variation 8 (table III) and to a slightly lesser degree for variation 2 (table III), suggesting in both cases a network with an NADP coupled cytosolic acetaldehyde dehydrogenase. Networks using an NAD dependent acetaldehyde dehydrogenase always resulted in a very high split ratio between the G6P dehydrogenase and isomerase, contradicting the results of the $^{13}$C-based MFA. However, in both cases the calculated contribution of the pyruvate decarboxylase is too high compared to the results of the $^{13}$C-based MFA. This can be the result of a minor error in the measurements, as even the same network except the pyruvate decarboxylase (variations 1 & 7) are still statistically acceptable. A lower contribution of the pyruvate decarboxylase in variation 2 would also change the split ratio at the G6P branch point to value comparable to the one obtained from
the $^{13}$C-labelling studies and thus eliminate this apparent discrepancy.

**tpiΔ mutant**

The *tpiΔ* mutant reached a steady state biomass concentration of 0.94 g/l on a 10:1 glucose and ethanol mixture, yielding 9.7 g biomass per Cmol substrate consumed. Significant amounts of glycerol, on average 0.8 g/l, was measured in the supernatant during steady state growth. The residual glucose concentration remained below 0.6 mM. Hence, the chemostat cultivation of the *tpiΔ* mutant resulted in the production of about 0.55 mol glycerol per mol glucose consumed. The higher biomass specific uptake rate of glucose and ethanol can only be partly explained by the glycerol production. The biomass yield of 12.3 g/Cmol$_{GLUC-ETOH-GOH}$ is significantly lower than the wild type’s 16.3 g/Cmol. Further reconciled rates are given in table II, generally confirming the results of Compagno et al. (Compagno et al., 2001). However, their large residual glucose of 11.7 mM was not observed under our growth conditions.

The exchange rates obtained from this chemostat cultivation were successfully balanced using a black box approach, statistical acceptance was also reached using the metabolic network II restrictions without cofactor balances, here a $X^2$-test value of 7.6 was obtained for a system with six degrees of freedom, compared to the cut-off value of 12.6. However, inclusion of the NADP or the ATP cofactor balances led to $X^2$-test value of 59.6 compared to the acceptable value of 15.5, and thus to the rejection of the data set for the mutant.

**metabolic fluxes**

<table>
<thead>
<tr>
<th></th>
<th>G6P isomerase</th>
<th>G6P dehydrogenase</th>
<th>MGLYOX bypass</th>
<th>ATP excess</th>
<th>NADP excess</th>
</tr>
</thead>
<tbody>
<tr>
<td>network II w/o cofactor balances$^1$</td>
<td>0.179</td>
<td>0.657</td>
<td>-</td>
<td>0.912</td>
<td>0.800</td>
</tr>
<tr>
<td>with methylglyoxal bypass$^2$</td>
<td>0.258</td>
<td>0.579</td>
<td>0.133</td>
<td>0.832</td>
<td>-</td>
</tr>
<tr>
<td>with methylglyoxal bypass II$^2$</td>
<td>0.358</td>
<td>0.479</td>
<td>0.100</td>
<td>0.852</td>
<td>-</td>
</tr>
<tr>
<td>with GOH dehydrogenase</td>
<td>0.568</td>
<td>0.271</td>
<td>-</td>
<td>0.695</td>
<td>-</td>
</tr>
</tbody>
</table>

$^1$ using NADP dep. acetaldehyde dehydrogenase & pyruvate dehydrogenase
$^2$ using glyoxylate
$^3$ using aldose reductase

**Table V:** reconciled metabolic fluxes of the *tpiΔ* mutant for different network II based variations without ATP balance. Statistical acceptance was accomplished only for the last three networks.

Analysis of the fluxes obtained from metabolic network II without the use of the cofactor balances for ATP and NADP showed a significant excess for NADPH and ATP (table V). The imbalance of NADPH was due to the high flux forced through the pentose phosphate pathway in order to match the measured glycerol production of less than one mole per mole glucose and send the remaining hexoses towards pyruvate. Subsequently, three different model variations were tested to overcome this imbalance. Firstly, the possibility of a further NADPH sink was investigated. A conceivable solution was a higher lipid content in the biomass. Secondly, alternative pathways for the further metabolization of DHAP would lessen the flux through the pentose phosphate pathway. Thus, the degradation of methylglyoxal formed from DHAP was investigated. Here, two possible routes were identified in yeast: a direct conversion via the methylglyoxylase (MG bypass.
I) or via the aldose reductase and the reductase and dehydrogenase for lactaldehyde to lactate (MG bypass II) (figure 4). The bypass pathway II allows a higher contribution of the pentose phosphate pathway and thus a slightly higher ATP yield from the metabolised glucose, due to its consumption of NADPH. The third and last variation degraded the excess DHAP via the dihydroxyacetone kinase and an NADP dependent glycerol dehydrogenase, thus simultaneously removing the NADPH formed in the pentose phosphate pathway.

![Diagram of metabolic pathways](image)

**Figure 4:** possible routes around dihydroxyacetone phosphate (DHAP) towards glycerol and pyruvate, the produced cofactors NAD, NADH, and NADP are shown next to the arrow of the reaction.

The first option, namely variations in the biomass composition could not explain the misfit between model and measurements. A higher lipid content of the biomass was insufficient to alleviate the NADPH problem, neither were changes to the carbohydrate or the protein content.

The second and third options were tested with the same twelve networks employed for the flux calculation of the wild type. The results for the resulting 48 variations are given in table VII. In contrast to the variation of the biomass composition, these options were able to close the NADP balance for certain networks. Of the 48 tested networks, 20 were statistically accepted, one of which exhibited a negative flow through the G6P dehydrogenase and must therefore be excluded. In any case, one of the methylglyoxal bypasses or the glycerol dehydrogenase (NADP) was required. Furthermore, statistical acceptance relied on the presence of pyruvate decarboxylase. Only then, the excess of 0.912 mol ATP/mol glucose found in the initial analysis could be eliminated, causing the majority of pyruvate to be metabolised via the decarboxylase. The additional energy
consumption of the methylglyoxal bypass is not sufficient on its own as can be seen from table V.

If included, the glyoxylase operated at 0.11 - 0.14 mol/mol glucose compared to the flux of 0.73 mol/mol through the lower glycolysis. Thus, around 17% of the pyruvate would be produced via this pathway, a substantial increase compared to the wild type, where Martins et al. (Martins et al., 2001) measured a relative flux of 1%. Very similarly, between 0.08 and 0.18 mol methylglyoxal were metabolised per mol glucose if modelled with the aldose reductase. Of all three options, the pathway via the glyoxylase resulted in the lowest split ratio between the G6P dehydrogenase and its isomerase, the highest ratio is obtained from the network employing the glyceraldehyde dehydrogenase (NADP).

NMR analysis was also performed for the tpiΔ mutant. Biomass harvested 35 hours after switching to 13C-labeled substrate was analysed for 43 spectra of various metabolic endproducts. Trehalose and levulinic acid could not be measured in the biomass. Again, the set of data was corrected to a label distribution at infinite time. Analysis of the different spectra confirmed the findings concerning the biosynthetic pathways of the wild type. The serine and glycine spectra corresponded better for the biomass of the mutant.

For the tpiΔ mutant the measurement set contained 139 relative intensities for 43 spectra,

![Figure 5: parity plot for the tpiΔ mutant of S. cerevisiae CEN.PK113-7D with the simulated relative intensities plotted over the measured relative intensities](image)

i.e. 96 independent data points to fit 18 degrees of freedom between the fluxes. The metabolic flux calculation resulted in a residual covariance-weighted residual between measured and simulated spectra of 1890. As was the case for the wild type, this minimized squared error is again larger than the statistically expected value of 76 (96 independent data points minus 18 parameters). The ratio between the found and the expected residual values is, however, considerably smaller in this case: 25 versus 92 for the wild type. This
indicates that the data are better fitted in this case, which is also reflected in the parity plot (figure 5).

![Figure 6: disparity plot of the measured relative intensities of S. cerevisiae CEN.PK113-7D and those of its isogenic tpi/Δ mutant](image)

The measured relative intensities of the same carbon atoms differed quite a lot between the wild type and the mutant strain. This is illustrated by the disparity plot in figure 6, where it can be seen that the difference between the measured data for the two strains is much larger than the misfit of the measured data by the simulated data (figures 3 and 5). This enables the clear discrimination between the flux patterns in both strains.

Compared to the wild type, the split ratio between G6P dehydrogenase and isomerase reverses, 0.48 mol flow via the pentose phosphate pathway against 0.36 mol/mol glucose through the glycolysis (figure 7). The methylglyoxal pathway is calculated as 0.07. The mitochondrial exchange flux of pyruvate is five times larger than in the wild type, but the exchange flux of C₄-TCA cycle intermediates is zero. Somewhat surprising is the outcome around the pyruvate branchpoint, were the contribution of pyruvate dehydrogenase was estimated to be zero, forcing all pyruvate thought the decarboxylase. A sensitivity analysis revealed a strong dependency of this ratio on the lysine-α spectra, which was not available for the mutant. Trials on the wild type data predicted a zero flux through the dehydrogenase of the wild type, when this spectra was left out, thus this result has to be treated with caution. Most other fluxes are comparable to those estimated for the wild type.

As was already discussed for the wild type, the confidence intervals of the fluxes in figure 7 cannot be calculated from the estimated measurement errors. Similarly to the wild type case we could, however, determine intervals within which the fluxes may be varied without significantly affecting the fit between the simulated and measured data set and without violating the constraint that all fluxes should have positive values. These intervals (table VI) again result from the observation that six of the eighteen singular values of the matrix that gives the sensitivities of the optimally fitted parameters to the ¹³C-data were

142
considerably (one >1700-fold, the rest >130000-fold) smaller than the other twelve. The table shows that, similarly to the wild type, the fluxes in the non-oxidative branch of the pentose phosphate pathway are badly determined. In contrast to the wild type, the pyruvate exchange over the mitochondrial membrane and the succinate scrambling rate in the mutant cannot be freely varied without affecting the fit.

Figure 7: flux pattern as determined by the $^{13}$C-labelling based MFA for the tpi1Δ mutant of S. cerevisiae CEN.PK113-7D grown at $D = 0.1$ hr$^{-1}$ in carbon limited chemostat culture using a network with methylglyoxal bypass, bidirectional fluxes in the metabolic network are indicated by double-sided arrows of which the closed head points in the direction of the net flux.

Comparison between the $^{13}$C-derived flux patterns and the ones obtained from the co-factor based MFA revealed, that for both networks I and II only the NADP dependent glycerol dehydrogenase can produce a split ratio at the G6P branch point larger than one, as found with the NMR-analysis. Furthermore, the NAD dependent acetaldehyde dehydrogenase can likely be excluded as possible option for network I. Variation 8 (table III), which was
found to provide a good match for the wild type, provided again the best match between both methods in the form of 8.4 (table VII). Variation 2, which also provided a good fit for the wild type, was unable to provide a reasonable fit between the cofactor and $^{13}$C-labelling based MFA results. A good fit is further found for variation 9.4, which lacks the pyruvate dehydrogenase, and variations 6.3 and 6.4, which are based on network I.

<table>
<thead>
<tr>
<th>reaction from to</th>
<th>flux minimum</th>
<th>maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 p5p &gt; s7p + g3p</td>
<td>0.12</td>
<td>0.34</td>
</tr>
<tr>
<td>s7p + g3p &gt; f6p + e4p</td>
<td>0.27</td>
<td>0.3</td>
</tr>
<tr>
<td>f6p + e4p &gt; s7p + g3p</td>
<td>0.11</td>
<td>0.15</td>
</tr>
<tr>
<td>p5p + e4p &gt; f6p + g3p</td>
<td>0</td>
<td>0.22</td>
</tr>
<tr>
<td>s7p + e4p &gt; f6p + p5p</td>
<td>0</td>
<td>0.22</td>
</tr>
<tr>
<td>f6p + p5p &gt; s7p + e4p</td>
<td>0</td>
<td>0.04</td>
</tr>
<tr>
<td>f6p + e4p &gt; e4p + f6p</td>
<td>0</td>
<td>0.04</td>
</tr>
<tr>
<td>s7p + p5p &gt; p5p + s7p</td>
<td>0.19</td>
<td>0.6</td>
</tr>
<tr>
<td>f6p + g3p &gt; g3p + f6p</td>
<td>0</td>
<td>3.38</td>
</tr>
<tr>
<td>s7p + e4p &gt; e4p + s7p</td>
<td>0</td>
<td>13.47</td>
</tr>
</tbody>
</table>

**Table VI:** error margins for the most insensitive metabolic fluxes determined with $^{13}$C-labelling in the tpiΔ strain, showing those fluxes that have absolute values larger than 0.005 and that can deviate more than 5% from their optimally estimated values.

The variations, 8.4 and 9.4, are characterised by the absence of a methylglyoxylate bypass, thus in contradiction to the model used for the $^{13}$C-labelling based MFA. Flux calculations based on the NMR data using a network without a methylglyoxal bypass resulted in the flux distribution depicted in figure 8, which differs from the previous calculation by its higher flux through the pentose phosphate pathway. This network structure, however, resulted in a worse optimal fit between the simulated and measured $^{13}$C-labelling data than the network that included the methylglyoxal bypass. Whereas the number of independent data minus the number of parameters only increased from 78 to 79 (due to the reduction of the number of free parameters by one), the minimized sum of squared errors increased from 1886 to 2977, i.e. by 58%. These outcomes of the $^{13}$C-labeling simulations lend support to the assumption that the methylglyoxal bypass is active to some extent in the mutant strain. Based on this observation, the cofactor based MFA was performed for four new network alternatives that contained a methylglyoxal bypass as well as a NADP dependent glycerol dehydrogenase (table VII, variations 8.5, 8.6, 9.5 and 9.6).
Table VII: 52 tested metabolic network options for the tpiA mutant varying in the chosen network (I or II, inclusive energetic parameters), the cytosolic acetaldehyde dehydrogenase (NAD or NADP dependent), and the use of the pyruvate dehydrogenase and the pyruvate decarboxylase. To address the NADPH excess three pathways were tested: a methylglyoxal pathway using the glyoxylase (bypass I), a methylglyoxal pathway using the aldose reductase (bypass II), and a combination of dihydroxy acetone kinase and an NADP dependent glycerol dehydrogenase. Furthermore, the degree of freedom (df) together with the \(X^2\)-test values (h) indicate the statistical acceptance (pass) of the measurement. The cutoff values for the 95% confidence interval are 15.5, 14.1 and 12.6 for a degree of freedom of 8, 7 and 6, respectively. Metabolic fluxes in moles per mol glucose are given for the main substrates, biomass, through glycolysis and pentose phosphate pathways, through the methylglyoxal pathway and around the pyruvate branch point.
Variations 8.5 and 8.6 resulted in a small negative flux through the irreversible pyruvate dehydrogenase, the setting of a boundary condition would thus lead to a zero flux at this point, making the variation 8 indistinguishable from variation 9. Results for variations 9.5 and 9.6 exhibited a small flux through the methylglyoxylate bypass (either I or II) of approximately similar magnitudes as the calculated by the $^{13}$C-labelling method, but had otherwise flux patterns very similar to variation 9.4.

\[\text{Figure 8: flux pattern as determined by the }^{13}\text{C-labelling based MFA for the }tpi1\Delta\text{ mutant of }S.\ cerevisiae\text{ CEN.PK113-7D grown at }D = 0.1\text{hr}^{-1}\text{ in carbon limited chemostat culture using a network without a methylglyoxal bypass, bidirectional fluxes in the metabolic network are indicated by double-sided arrows of which the closed head points in the direction of the net flux.}\]
Discussion and Conclusions

Wild type

The good fit of the metabolic network models for the wild type data confirmed the validity of the approach for the MFA using cofactor balances, even though the energetic parameters were obtained from chemostat experiments using either glucose or ethanol as substrate. Interpolation of these findings to mixed substrate chemostat experiments are thus valid, thus confirming the findings of Vanrolleghem et al. (Vanrolleghem et al., 1996) also for the new compartmented network model.

The plain analysis of the $^{13}$C-spectra revealed important information about the source of the metabolic precursors. It thus allowed the verification of some essential network assumptions and enabled an educated guess about the localisation of biosynthetic pathways. Using these insights the reversibility of the hydroxymethyl transferase for the formation of glycine was confirmed. Also based on this analysis was the placement of the alanine aminotransferase in the mitochondria, confirming the assumptions of Maahieimo et al. (Maahieimo et al., 2001). The same confirmation for metabolic routes was obtained from the measurements on the biomass of the $tpi\Delta$ mutant.

The split ratio between G6P dehydrogenase and isomerase was estimated to be 0.32:0.42 from the results of the $^{13}$C-labelling based MFA. The results are somewhat lower than those of Christensen et al. (Christensen et al., 2001) and Gombert et al. (Gombert et al., 2001) who calculated a flux of respectively 0.43 and 0.44 of the initially consumed glucose through the pentose phosphate pathway for chemostat growth of the same strain on pure glucose. This is probably due to the use of the NADP dependent acetaldehyde dehydrogenase for the consumption of parts of the co-fed ethanol in our experiments. This was confirmed by the co-factor based MFA, where only this option was able to reproduce results similar to the ones of the $^{13}$C-labelling based MFA.

Somewhat surprising was the flux distribution through the pyruvate branch point. The results of the NMR study suggest, that almost 1/3 of the pyruvate is metabolised via the pyruvate bypass. This was mirrored by the better fit for the co-factor based MFA with networks including the pyruvate decarboxylase, which frequently lead to a substantial flux through the pyruvate decarboxylase. These results are in contradiction to the commonly applied maximal yield principle, which leads to an expectedly absent flux through the energetically less favourable pyruvate bypass. A significant flux through the pyruvate dehydrogenase was also found by Gombert et al. (Gombert et al., 2001), who calculated a ratio between decarboxylase and dehydrogenase of 0.45:0.40 for similar growth conditions under pure glucose limitations.

Comparing the results from the co-factor based MFA with the values obtained using the $^{13}$C-labelling technique, most network options can be confidently excluded. Clearly, the use of an NAD dependent acetaldehyde dehydrogenase can not be brought into agreement with both modelling approaches. The operation of both, the pyruvate dehydrogenase and
the decarboxylase, reduces the number of likely metabolic networks to only two, variations 2 and 8 in table III.

**tpiΔ mutant**

Initially, the results of the tpiΔ mutant could only be partly explained using the cofactor based MFA. Obviously, the restrictions of the NADP balance could not be fulfilled by the model used for the wild type without the triose-phosphate isomerase reaction and extended by an active glycerol export system. The problem could be traced to the split ratio between the flux of the G6P dehydrogenase and the isomerase. This split ratio directly influences the expected amount of excreted glycerol, as a higher flux through the pentose phosphate pathway will lead to a decrease in flux towards DHAP and subsequently to glycerol. Due to the low yield on glucose, the deletion mutant requires much less NADPH relative to the glucose uptake than the wild type. However, the glycerol measurements would indicate a split ratio comparable to the latter, as can easily be calculated from the glycerol yield on glucose. If implemented, this would lead to a large surplus of NADPH.

Several solutions are conceivable. The first option would allow a lower split ratio and would require the further metabolism of the formed DHAP through, e.g. an additional methylglyoxal bypass and thus decrease the required glycerol excretion in accordance with the measured values. An alternative solution would involve a higher split ratio and an additional sink for the NADPH. This would be possible through the adaptation of the biomass composition or an NADP coupled reaction in the glycerol production pathway. Changes in biomass composition, e.g. a higher lipid content were however insufficient for this purpose. Hence, the NADPH dependent glycerol dehydrogenase remained the only possible option. In all three cases, with both methylglyoxal bypasses or the glycerol dehydrogenase, the NADP balance could be closed and statistical acceptance be reached for twenty different metabolic network variations (table VII).

A large flux through pyruvate decarboxylase is required to close the ATP balance of the tpiΔ mutant when analysed with the cofactor based MFA. A comparison of the energetic costs of the pentose phosphate pathway versus the glycolysis as an alternative energy sink shows, that the disadvantage of the former is relatively small for aerobically grown cells. Variations in the split ratio would therefore be unable to explain the higher ATP requirement of the mutant. Other options to utilize the calculated ATP excess also proved to be insufficient, e.g. the contribution of the flux through the methylglyoxal bypass I or II is too small to be directly linked to a possible ATP sink. The flux through the pyruvate decarboxylase is confirmed by the 13C-labelling based analysis, which predicts an insignificant role for pyruvate dehydrogenase in the mutant's metabolism during growth on a mixture of glucose and ethanol. Both modelling approaches agreed very well in these findings which gives additional confidence in the results. However, the lack of the lysine-α spectra in the mutant data set makes these observation less reliable than those of the wild type.

Co-factor based MFA proved to be capable of eliminating half of the considered network options for the tpiΔ mutant. While acceptance levels and predicted fluxes were very similar for all accepted cases, the main differences were found in the split ratios between the G6P dehydrogenase and the G6P isomerase. Therefore, narrowing down the remaining
options required the use of the $^{13}$C-labelling based MFA. The results of the NMR study suggest a split ratio above one, which makes the methylglyoxylase bypass I (variations x.2) an unlikely option leaving twelve variations. Good fits were reached for variations 8, 9 and 6, but not for variation 2.

Under the assumption that the wild type strain employs the same metabolic network as the mutant with the exception of the deleted triose-phosphate isomerase, network II with an NADP dependent acetaldehyde dehydrogenase is the most likely configuration. None of the other variations simultaneously fits the wild type and the mutant data. Very likely, the mutant additionally uses the NADP dependent glycerol dehydrogenase with only a small contribution of the methylglyoxylate bypass. However, the operation of the methylglyoxal bypass can neither be excluded nor confirmed with absolute certainty. The ethanol requirement for the tpiΔ mutant could not be explained with this model, and must, under the assumption of a correct metabolic network be attributed to kinetic limitations.

The co-factor based MFA for both strains relied on a set of energetic parameters describing the aerobic ATP regeneration, and which were obtained using the maximum yield principle, thus not including a flux through the pyruvate decarboxylase. Its inclusion would likely raise the $P/O$ ratio or lower the $k_{\text{ATP}}$. Therefore, one could not expect a lower ATP production in the cases here presented should the parameters have been determined with a network deviating from this assumption. A more detailed analysis of the energetic parameters would require the application of the $^{13}$C-labelling technique.

The minimization of the difference between the simulated and measured $^{13}$C-labelling data of both the wild type and tpiΔ mutant strain resulted in a good fit with respect to the absolute deviations. There is an absence of large outliers that could indicate missing or faulty reactions. However, when the minimized deviations between the simulated and measured data are weighted by the measurement error that were estimated from the NMR noise it is found that the fits are statistically unacceptable. As it is not unthinkable that the NMR measurement noise is not the only source of errors, it is recommended to generate multiple measurement data sets to check whether the NMR noise really covers the variation in the data. If this is the case, the conclusion must be that the biochemical model is incomplete and needs further refinement in order to determine the true metabolic fluxes.

The presented results demonstrate the capacity of both modelling approaches to not only quantify metabolic fluxes in metabolic networks, but also to test the validity of a network. While the co-factor based MFA is capable of relatively easily determining flux patterns for a large number of different networks, the more labourious and computationally demanding $^{13}$C-labelling approach is capable of determining parallel pathways unaccessible to the first method. Analysis of the NMR spectra enables a distinction between certain pathway options, as seen in the example of the localisation of the alanine aminotransferase. However, only the combination of these mathematical tools with the study of the wild type strain and its isogenic deletion mutant allowed for the final reduction in the number of possible network options.
Acknowledgements

We would like to acknowledge Diana Ros Riu, who helped with the spectra analysis, and Wouter Berendse and Ko Vinke for conducting the fermentations. Danilo Porro from the Università degli Studi di Milano-Bicocca kindly provided the tpi1Δ-mutant strain. The project was partly funded through the European Framework IV program and by the TUDelft DIOC-programme ‘Mastering the Molecules in Manufacturing’.

References


cerevisiae through metabolic network modelling verified by chemostat cultivation.


Chapter 7

Rapid sampling for intracellular metabolite determination

published as

Improved rapid sampling for in-vivo kinetics of intracellular metabolites in _Saccharomyces cerevisiae_


in

Biotechnology and Bioengineering
Vol. 75(4), 406-415, 2001
Rapid sampling for intracellular metabolite determination

Abstract

An integrated approach is used to develop a rapid sampling strategy for the quantitative analysis of the in-vivo kinetic behaviour based on measured concentrations of intracellular metabolites in \textit{Saccharomyces cerevisiae}. Emphasis is laid on small sample sizes during sampling and analysis. Sub-second residence times are accomplished by minimizing the dead volume of the sterile sampling system and maximise flow rates by applying vacuum to the sampling tubes in addition to the over-pressure in the fermentor. A specially designed sample tube adapter facilitates sampling intervals of 4 to 5 seconds for various test tube types. Statistical analysis of the results obtained from enzymatic and liquid chromatography coupled to tandem mass spectrometry (LC-MSMS) analysis of the metabolite concentrations was used to optimize the sampling protocol. The most notable improvement is reached through the introduction of vacuum drying of the cell extract. The presented system is capable of reliably dealing with fermentor samples as small as one gram with a variation of less than 3% and is thus ideally suited for intracellular measurements on small lab-scale fermentors.
Chapter 7

Introduction

Measurements of intracellular concentrations were initially driven by the desire to obtain a better understanding of the in-vivo enzyme kinetics (Gancedo and Gancedo, 1973) and enzyme expression patterns (Entian et al., 1977). Rapid development of genetical tools and computer capacities shifted the interest towards the complete biological system with the intention of specific improvements of the microorganism through genetic alterations (Bailey, 1998). Today’s DNA technologies allow the directed reprogramming of the microbial metabolism towards an increased production of a desired product. With that, an enhanced activity of the involved pathways might ultimately lead to shortages of the metabolic precursors provided by the central metabolism. Therefore, the complex interdependencies of the latter have to be well understood to overcome these limitations and allow consistently successful metabolic reprogramming. Further insight into this system can be obtained with mathematical models describing the in-vivo kinetics of the reaction network of the primary metabolism, which will be based on measured intracellular concentration profiles.

In a first step, a proposed kinetic model of the metabolic system is tested against controlled perturbations of a metabolic reference state $f_0$ with defined concentrations of metabolites $x_e$, enzyme levels $e_v$, and reaction rates $v_v$. Perturbations can be induced by changes of the extracellular concentrations through the discrete addition of a component or a change of their flow rates, causing fluctuations in the intracellular reaction rates $v$ and thus the metabolite levels $x$, followed by changes in enzyme activities $e$. The system will either return to its original reference state $f_0$ or reach a new steady state $f'$, if no further transients are forced upon the system. Of particular interest is the time window of up to 300 seconds in which enzyme concentrations remain practically unchanged and their levels $e$ can be considered constant ($e = e_v$) as a first approximation. The variables in this period are thus the reaction rates and the intracellular concentrations of the metabolites, of which only the latter ones are measurable (Visser et al., 2000). Thus, the reaction rates can only be deduced from the metabolite concentrations using dynamic metabolite balances.

Generally, the measurement of intracellular concentrations proceeds via fast sampling combined with a quenching step, which ensures the rapid inactivation of the metabolism and freezes the cell’s state at the moment of sampling. Cell extraction, sometimes preceded by the separation of the extracellular fermentation medium, makes the metabolites accessible to the different analytical methods employed (Rizzi et al., 1997; Weuster-Botz, 1999). The details for each step have to be adapted to the particular organism under investigation (Jensen et al., 1999), the here presented protocol was tailored towards the yeast Saccharomyces cerevisiae.

For many components of interest, intracellular turnover rates are in the order of one second, e.g. ATP, thus intracellular measurements require sampling with much shorter residence times and instant deactivation of the enzymatic activity. The maximal duration of the complete sampling procedure is suggested to be in the order of a few hundred milliseconds (Weuster-Botz and de Graaf, 1996). Hence, considerable effort has been put
into the development of a fast sampling and inactivation procedure. Theobald et al. (Theobald et al., 1993), Iversen (Iversen, 1981), and Larsson et al. (Larsson and Törnkvist, 1996) used sterilizable miniature valves with minimal dead volume to reduce the residence time to approximately 100 milliseconds. Sampling intervals of several seconds are possible with these setups. Higher time resolutions require the use of an incubation tube in which the progress of a dynamic response is determined by the length of the tubing (Weusten-Boitz, 1999).

A fast metabolic arrest is commonly achieved through rapid changes in temperature or pH accomplished by mixing the sample with an appropriate quenching solution, e.g. liquid nitrogen (Hajjaj et al., 1998), liquid CO₂, cold methanol (de Koning and van Dam, 1999; Gonzalez et al., 1997; Hajjaj et al., 1998; Schäfer et al., 1999; Smits et al., 1998; Smits et al., 1996; Theobald et al., 1997), boiling ethanol (Ciriacy and Breitenbach, 1979; Gonzalez et al., 1997), cold perchloric acid (Cole et al., 1967; Franco et al., 1984; Larsson and Törnkvist, 1996; Theobald et al., 1997), or trichloroacetate (Kopperschläger and Augustin, 1967). The exact treatment has to be adapted to the components of interest. Cell integrity has to be maintained during quenching, if the separation of intra- and extracellular fluids is required. Enzyme based analysis of the metabolites in the extract requires the removal of any organic solvent, which can be accomplished through drying e.g. by film evaporation (Gonzalez et al., 1997). Furthermore, the dried samples can easily be stored for later analysis without fear of degradation of the components of interest.

The presented methods led to valuable information on intracellular metabolite concentrations of S. cerevisiae and its response to external disturbances (Rizzi et al., 1997; Schäfer et al., 1999), with the cold methanol method being preferred in yeast research, because of its ability to distinguish between intra- and extracellular concentrations and its possibility to concentrate the metabolites. Nevertheless, the described systems suffer from several shortcomings, most notably on the sampling size. Large amounts of broth withdrawn during sampling will cause significant disturbances of the fermentor operation, especially in small lab-scale setups. Control of the sampled amount is usually accomplished by a combination of over pressure in the fermentor and vacuum in the sample tube, making a reproducible sample size difficult (Theobald et al., 1997). However, a consistent and precise sample size is essential to ensure the exact ratio between quenching solution and sample and thus to achieve a consistent and fast temperature drop while preventing cell wall rupture through the formation of ice crystals. Care also has to be taken during the following washing steps to keep the temperature below -20°C. Furthermore, accuracy and effectiveness of washing and boiling are rarely reported. These problems become even more eminent for smaller sample sizes, because of the larger influence of wall effects and their increased temperature-sensitivity due to the small heat capacity of each sample.

This article presents an integrated approach to develop an advanced rapid sampling strategy for the quantitative analysis of intracellular metabolite concentrations in Saccharomyces cerevisiae. Emphasis is laid on the use of small sample volumes, allowing investigation on lab-scale fermentors, while maintaining a precise ratio between quenching liquid and sample and improving the overall accuracy. The proposed system allows the fast and reproducible handling of large numbers of samples here used to study the
reproducibility of the sampling, the sample treatment, and the analysis. Development and improvement of the procedure was driven by statistical analysis of the obtained data.

Materials and methods

Biomass was obtained through carbon-limited, aerobic chemostat cultivation of *Saccharomyces cerevisiae* CEN.PK113-7D on doubled mineral medium (Verduyn et al., 1992) with 27.1 g/l glucose and 1.42 g/l ethanol mixture in a 7-litre fermentor (Applikon, Schiedam, The Netherlands) to obtain approximately 15 g biomass (dry weight) per litre. A growth rate of 0.05 hr⁻¹ was maintained for a culture volume of 4 litres at a pH of 5.0 and at 30°C. Over pressure of 0.3 bar and a stirrer speed of 700 rpm ensured a dissolved oxygen content of never less than 30%.

Total organic carbon (TOC) of broth and filtrate was measured on a 5050A TOC-analyser (Shimadzu Corporation, Australia) using potassium hydrogen phthalate as standard. Temperatures were measured using a iron-constantan thermocouple calibrated against a water/ice mixture. The small mass minimized the influence of the probe on the sample temperature.

The quenching solution consisted of 60 w/w % analytical grade methanol (ACROS, Geel, Belgium) and water purified in a milli-Q UFplus system (Millipore, Bedford, MA, USA), buffered with 10 mM HEPES (Merck, Darmstadt, Germany) and adjusted to pH 7.5 with 3 M KOH (Baker, Deventer, The Netherlands). The extraction solution contained 75 w/w % ethanol in water, buffered with 10 mM HEPES and adjusted at pH 7.5 as the quenching solution. No buffers were used for TOC and liquid chromatography - mass spectrometry (LC-MSMS) measurements.

Metabolite analysis was performed enzymatically for NAD, NADP, and ATP based on the Bergmeyer protocols. Dried samples obtained from 1 gram of broth were dissolved in 0.5 ml demineralized water prior to analysis. NAD and NADH were measured using enzymatic cycling on 96-well microtitre plates. NAD was measured in a 200 µl sample volume after destruction of NADH by addition of 30 µl HCl (0.1 M) and neutralization by 30 µl NaOH (0.1 M). NADH was measured following NAD destruction through addition of 30 µl NaOH (0.1 M), heat treatment at 60°C for 10 min, and neutralisation. 50 µl of pretreated sample was buffered in 200 µl 50 mM Tris & 5 mM EDTA solution at pH 7.8. Aliquots of 10 µl of ethanol (100%), PES solution (10 mg/ml), MTT solution (3 mg/ml), and alcohol dehydrogenase solution (30 U/ml) were added. Absorbance was measured for 20 min at 590 nm in a HTS7000 Plus BioAssay reader (Perkin Elmer, Norwalk, CT, USA). NADP and NADPH measurements followed the same protocol with ethanol being replaced by a 20 mM solution of G6P and the reaction being catalysed by G6P dehydrogenase (8 u/ml).

Glucose-6-phosphate (G6P) was analysed enzymatically following the Bergmeyer protocol adapted for 1 ml quartz cuvettes. 800 µl of TEA buffer containing 27 µM NADP was mixed with10-50 µl sample. The reaction was allowed to proceed for 3 minutes after the
addition of 5 µl of G6P-dehydrogenase (5 U/min) buffered in 3.2 M ammonium sulfate. Absorption of NADPH was measured at 339 nm with an HP8453 diode array spectrophotometer (Agilent, Palo Alto, CA, USA).

ATP analysis was performed on 50 µl aliquots using an ATP Bioluminescence Assay Kit CLSII (Boehringer, Mannheim, Germany) on a Costar 96-well micro-titre plate in a mediators PhL plate reader (mediators diagnostic systems, Vienna, Austria). Samples of the resuspended intracellular metabolites were diluted 500 fold, the supernatant was analysed 400-fold diluted. All chemicals were purchased from Boehringer.

The glycolytic intermediates glucose-6-phosphate (G6P), glucose-1-phosphate (G1P), fructose-6-phosphate (F6P), fructose-1,6-bisphosphate (F16P), 2- and 3-phosphoglycerate (2PG, 3PG), phosphoenol-pyruvate (PEP), and pyruvate (PYR) were redissolved in 500 µl MilliQ water and analysed simultaneously in 10 µl using high performance using LC-MSMS. High-performance anion exchange chromatography was performed with an Alliance pump system (Waters, Millford, USA) using a sodium hydroxide gradient on an IonPac AS11 (Dionex, Sunnyvale, USA) with a AG11 pre-column (Dionex) at a flow rate of 1 ml/min coupled via an ion-suppressor (ASRS-Ultra, Dionex) to a tandem mass spectrometer (Quattro-LC, Micromass, Cheshire, England). The mass spectrometer was operated in negative mode with a nebulizer gas flow of 75 l/hr (source block temperature of 80°C), desolvation gas flow of 680 l/hr at 250°C, and a capillary voltage of -2.7 kV generating ion fragments via collisionally induced dissociation with Argon at 5*10⁻⁴ mbar.

![Figure 1: rapid sampling port](image-url)
Rapid sampling procedure

Reproducible withdrawal of small sample sizes for intracellular measurements drawn from lab-scale fermentors demands an adapted sampling strategy to result in reliable and accurate data. Further requirements arise from the desire to distinguish between intra- and extracellular concentrations, which can be accomplished in yeast cells by quenching in cold methanol. This method allows separation of the cells from the fermentation broth through centrifugation but requires sample handling at low temperatures to eliminate or at least minimize reaction rates before cell extraction and protein denaturation in boiling ethanol (Entian et al., 1977; Gonzalez et al., 1997) or chloroform (de Koning and van Dam, 1999). The presented procedure is based on previously reported protocols, but overcomes several of their drawbacks.

Sampling

One of the most popular methods was developed by the group of Reuss (Theobald et al., 1997). Their sterilizable micro valve has the lowest dead volume of all presented ports,

**Figure 2**: scheme of rapid sampling setup

operation is straightforward, and it has a history of successful application. Samples with a typical size of 3 ml are obtained within a closed vessel, which is an important aspect when dealing with genetically modified organisms. Nevertheless, the method has a number of disadvantages. The dead volume of 200 µl containing stagnant broth has to be further minimized to be suitable for sample sizes of around 1 ml. Test tubes have to be sealed and evacuated prior to sampling. Injecting the sample by means of a needle into the evacuated
and sealed test tube is susceptible to blockage and bending of the needle, premature loss of vacuum with a subsequent deviation of the sample size, and bad mixing due to tilted sample tubes. Rapid succession of samples is further inhibited by the danger posed by the sharp needle to the operator. We further experienced sterility problems with the valve seat. The following alternative was developed to overcome these difficulties.

The modified system consists of a submerged capillary port with an inner diameter of 1 mm and a length of 80 mm, placed inside a stainless steel cylinder to fit a standard fermentor port (figure 1). Silicon tubing (ID 0.8 mm) connects the port via a Y-piece to a waste container and the sample tube adapter (figure 2). A pinch valve directs the flow to either of them, switching times are controlled electronically through a mains-controlled digital counter designed and constructed in-house. During periods with no sampling, the silicon tube is clamped between the port and the Y-piece, while the two other ends are flushed and filled with 70% ethanol and connected to each other to form a closed loop.

The tube adapter closes the top of any standard sized test tube airtight with a foam pad against which the tube is pushed. Two plates placed perpendicular to each other guide the tube during manual changing and ensure its straight positioning. Two stainless steel tubes lead through the foam closure into the test tube (figure 3). During sampling the smaller, centrally placed one (ID 0.8 mm) is connected to the silicon tube coming from the fermentor. The second one (ID 1.4 mm) is used to evacuate the tube prior to sampling; a silicon pump tubing (Masterflex 96400 x 14) leads via a T-piece to a 2-litre vessel, which is kept at a constant vacuum of approximately 100 mbar; the other end is kept open. A second electronically controlled pinch valve enables switching between opening to ambient pressure and the vacuum container (figure 2).

The test tubes are filled with 5 ml quenching solution, weighted, capped to prevent water vapour condensation in the cold solution, and placed into the cryostat at -40°C at least two hours before sampling. The tubes are uncapped just prior to the sampling, weighed afterwards to determine the sample size, and closed again until further handling. During the sampling operation
fermentation broth is constantly flowing at a flow rate of about 0.5 ml/sec into the waste container. After placing a tube filled with 5 ml of -40°C quenching solution under the tube adapter, the start of a 3-step valve operating sequence is triggered manually: As the first step, the pinch valve 2 (figure 2) opens the tube leading to the vacuum container, one second later the pinch valve 1 switches from waste container to tube adapter, and after a further interval of around 0.7 seconds both valves fall back into their starting position. During the first interval vacuum is drawn which keeps the tube firmly in place below the tube adapter, during the second the flow of broth increases to about 1.5 ml/sec due to the vacuum in the tube adding 1 ml of sample with a straight jet into the sample tube as suggested by Harrison and Maitra (Harrison and Maitra, 1969). After the third step the vacuum in the tube is released and the tube with the sample is quickly placed back into the -40°C cryostat after short vortexing. The total inner volume of the sample port, the tubing, and the tube adapter is about 100μl, of which only the 50 μl between the Y-piece and the orifice of the tube adapter contain a stagnant liquid during sampling.

Separation of cells and extracellular liquid

Biomass and supernatant can be separated prior to the cell extraction to distinguish between intra- and extracellular concentrations of metabolites such as pyruvate, lactate, or acetate. This requires a quenching step which preserves the cell integrity, which is not maintained with acidic or caustic quenching methods, where quenching and cell extraction proceed simultaneously. However, it can be realised for yeast with organic solvents, e.g. cold methanol, in which metabolic arrest is accomplished through a combination of solvent based enzyme inhibition and low temperature. Therefore, sample handling has to be performed at temperatures below -20°C to prevent any reaction from proceeding until complete protein denaturation during the extraction step.

Cell separation was accomplished through 5 min centrifugation of the sample tube in a stratus biofuge (Heraeus, Hanau, Germany) at 2000 rpm using a precooled rotor (-40°C), decanting of the supernatant, resuspension of the cell pellet in 5 ml cold quenching solution (-40°C), and a second, identical centrifugation and decanting step. The tubes containing the washed pellet were placed back at -40°C. The centrifugation speed of 2000 rpm and gentle braking (setting at “2”) provided a solid yeast pellet which could be resuspended speedily. Quick sample handling was essential, since temperature in the test tubes would increase to values above -20°C if handled slowly.

Cell extraction

Extraction of the metabolites in boiling ethanol solution was proposed first by Entian (Entian et al., 1977) as being suitable for a wider range of metabolites than pH-based extraction methods and less hazardous than the chloroform extraction. Sample tubes containing the cell pellet were taken directly from the -40°C cryostat bath, topped with 5 ml of boiling extraction solution, vortexed, placed for 3 min in a 90°C water bath as proposed by Gonzalez (Gonzalez et al., 1997), and finally cooled by placing them back into the -40°C cryostat.
Sample drying

Removal of the extraction liquid containing ethanol and traces of methanol from the quenching step is required for the analysis of the samples either enzymatically or with reversed phase HPLC. Initially, the extracted samples were dried in a film evaporator with a water bath temperature of 40°C using 50 ml round bottom flask following the protocol of Gonzalez (Gonzalez et al., 1997). Later samples were desiccated in their original tubes under vacuum at about 0.2 mbar in a speedvac centrifuge (Savant, NY, USA) with precooled rotor combined with a cold-trap and RW12 rotary vacuum pump (Edwards, England). This eliminated the transfer of the extracted cell solution from the test tube to the round bottom flasks and thus avoids any loss of sample. The drying process was followed through pressure measurements by two APG-M active Pirani gauges (Edwards, England) in the centrifuge dwell and at the pump inlet. The vacuum drying process was terminated when the pressure in the centrifuge well equalled the one at the pump inlet. At this point the samples were visibly dry. Tests with partial drying as done by de Koning (de Koning and van Dam, 1999) were abandoned, because of the difficulties to precisely determine the remaining volume. Sample drying was usually completed within about 12 hours. The samples stayed liquid throughout the whole drying process, the starting temperature was -40°C and only rose to about 10°C towards the end. Vacuum drying allowed simultaneous desiccation of a large number of samples. In our case the maximal number of samples processed simultaneously was limited to 22 by the rotor configuration. Dried samples were stored in the initial test tubes at -20°C until analysis.

The dried samples were redissolved by adding 500 µl water prior to analysis, cell debris was removed through centrifugation in a table centrifuge. Consequently, the metabolite concentrations doubled compared to the original sample, higher ratios up to about 10 are easily feasible to lower the detection limit of the metabolites, but were not needed in the current situation.

Results

Sampling

Sampling was performed only on fermentations which had reached steady state with a constant biomass concentration of about 14 g/l. Sterility of the culture was maintained for a period of more than 1400 hours, during which samples were taken at least once per week. Flow characteristics of the sampling valve stayed constant during the prolonged fermentation period. No plugging of the sample port was ever observed.

Sets of 12 samples from steady state conditions were usually taken within 70 seconds with an average interval of 6.5 seconds. Intervals of less than 5 seconds were possible. The opening time of valve 2 (figure 2) was adjusted to around 0.7 seconds during which exactly 1 ml of fermentation broth was drawn into the cold methanol solution. The high volume flow ensured a residence time in the sampling unit of below 100 milliseconds. Switching between tube adapter and waste reduced the stagnant liquid in the port to less than 50 µl. A high flow velocity at the sampling orifice facilitated a straight jet into the
quenching solution and thus good mixing of both liquids essential for an instant metabolic arrest. Harrison and Maitra (Harrison and Maitra, 1969) measured a mixing time of below 80 milliseconds for a similar setup; this was further confirmed by Iversen (Iversen, 1981). Previous attempts of spraying the broth into the quenching solution caused the formation of a layer of broth on top of the quenching liquid, thus never accomplished complete mixing of sample and quenching solution unless vortexed. Furthermore, direct contact between the cold tube's wall (-40°C) and the fermentation broth could not be excluded for the narrow tubes used, causing freezing of fine films in the absence of methanol leading to cell wall disruption or slower quenching for larger droplets.

**Error analysis**

Initial results of the metabolite measurements by LC-MSMS were used for statistical error analysis to quantify the uncertainties introduced by the different steps of the protocol. The data exhibited large deviations between the twelve different samples of a single set drawn from a steady state culture. Concentration measurements of the individual glycolytic components G6P, G1P, 2+3PG, and PEP had similar standard deviations of on average 15.5% between these samples. Normalisation to the average of each component revealed a clustering pattern of the set's results which was analysed to identify the main error sources (figure 4A). The scatter within each sample's cluster reflected the precision of the analytical method, its standard deviation of about 5.5% could thus be equalled with the precision of the LC-MSMS analysis. The 14.5% variation of the cluster's centres indicated the errors introduced by the preceding sampling procedure.

The error distribution was confirmed by the analysis of the nucleotides, which was also performed using film evaporation as drying method. Measurements of NAD varied with a standard deviation of 4.6% for a single sample and 13% between samples of one set. Small discrepancies between these values and the ones obtained by LC-MSMS analysis were very likely due to the relatively small number of measurements performed enzymatically. Error analysis was not performed for the other nucleotides for the same reason. The initial search for possible error sources was therefore concentrated on the sampling.

The sample size was highly independent of the initial placement of the tube and varied with a standard deviation of less than 2% thus maintaining the critical ratio of 1:5 between sample and quenching solution. This is a clear improvement compared to the previous setup, with which we experienced a standard variation of about 10% in the sample size.

Experience showed that during washing the temperature easily rose above -20°C for sample tubes containing about 6 ml liquid when using the -20°C setting of the centrifuge. Precooling the rotor to -40°C prevented this for both centrifugation steps, temperatures stayed below -30°C when the washing step was done with a quenching solution chilled to -40°C. Nevertheless, speedy handling during resuspension was required to prevent excessive warming of the sample and to keep the temperature below -20°C until the addition of boiling ethanol.
Figure 4: set of normalized LC-MSMS results, obtained with film evaporation (A) or with vacuum drying (B). Measurements belonging to one sample form distinct clusters. Variations between these clusters indicate the noise associated with sample treatment, variations within each cluster are due to variability in the analysis. Differences between (A) and (B) show the improvement gained by the vacuum drying.

The film evaporation together with the decanting step was identified as the main source of error. Replacing it by vacuum drying decreased the average standard deviation of the LC-MSMS measurements from 15.5% to 8%. The cluster centres of the normalized results were now located closer to the base line (figure 4B). While the variations within the clusters remained the same, the standard deviation of the cluster’s centres decreased to 5.5% confirming the positive affect of the improved sample handling.

**Quantification and recovery**

Quantification of intracellular concentrations might be distorted by premature cell lysis during sampling. Cell integrity was therefore verified by ATP analysis of the vacuum dried and redissolved supernatant. In this way, 19 μmol ATP/l was found in the supernatant
compared with 11 μmol/l in the filtrate of the fermentation broth (pore size 0.45μm, Gelman Science, Ann Arbor, MI, USA). The 8 μmol/l difference can be attributed to cell leakage during the sampling and quenching steps, if the possibility of ATP hydrolysis in the filtrate through extracellular ATPase is disregarded. This amount corresponds to about 10% disrupted cells. The loss of cell integrity during sampling and decanting was also estimated through TOC analysis in the supernatant and the filtrate. Of the 7500 mg/l of organic carbon in the broth, 370 mg/l are found in the filtrate. Thus, the maximum measured TOC in the supernatant of 1000 mg/l equals a 10% loss of biomass confirming the ATP based results.

The influence of boiling time on the metabolite extraction was tested for ATP, NAD, NADH, NADP, and NADPH. While ATP extraction was independent of the duration of ethanol boiling, the other nucleotides with the exception of NADP showed a concentration maximum at 3 minutes. NADP extraction reached its peak slightly later between 4 and 5 minutes. These results confirmed the findings of Gonzales et al. (Gonzalez et al., 1997). Hence, a standard boiling time of 3 minutes was used for all samples and the so gained NADP values had to be corrected with a factor of 1.15 to compensate for the incomplete extraction.

<table>
<thead>
<tr>
<th></th>
<th>G6P</th>
<th>G1P</th>
<th>F6P</th>
<th>F16P</th>
<th>2PG</th>
<th>PEP</th>
<th>PYR</th>
</tr>
</thead>
<tbody>
<tr>
<td>added (μmol)</td>
<td>50</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>50</td>
<td>50</td>
<td>2.5</td>
</tr>
<tr>
<td>recovery (%)</td>
<td>105</td>
<td>100</td>
<td>99</td>
<td>128</td>
<td>115</td>
<td>92</td>
<td>60</td>
</tr>
</tbody>
</table>

Table I: recovery of metabolites, metabolites were added in 100μl aqueous solution with concentrations close to the one found in the original sample. The solution was added after the washing step, before boiling. The recovery values are averages of three independent trials.

Further disturbances were introduced by a HEPES accumulation on the HPLC column, leading to increasing standard deviations. Furthermore, comparison between buffered and unbuffered samples revealed a 20% difference for G6P. Enzymatic measurements of G6P did not exhibit this difference and resulted in 2.2 μmol G6P/l. This corresponded very well with the results obtained by HPLC-MSMS for HEPES-free samples. Subsequently, HEPES was omitted for samples analysed by HPLC-MSMS.

Metabolite recovery was tested on three occasions through standard additions to the washed cell pellet. 100μl of a standard solution with 0.1 mM G1P, 0.5 mM G6P, 0.1 mM F6P, 0.1 mM F16P, 0.5 mM 2PG, 0.5 mM PEP, and 25μM PYR were added before the boiling step. Recoveries were calculated from the differences between untreated and spiked samples. G6P, G1P, F6P and PEP were recovered with 105, 100, 99, and 92% respectively (table 1). Recoveries for F16P and 2PG were with 128 and 115% slightly too high. In the case of 2PG, this might simply be due to its error margin. F16P was repeatedly found in higher concentrations than justified by its addition, the average recovery of 125% could not be explained to our satisfaction. Recovery of pyruvate was insufficient, about
half was lost during drying and analysis, but not as initially expected during the hot ethanol extraction.

**Chemostat results**

The intracellular concentration was measured periodically on a chemostat culture with constant media composition as described in a previous section. The average results measured enzymatically were 2.4, 0.4, 0.2, and 0.05 μmol/gDW for NAD, NADH, NADP, and NADPH respectively. The intracellular concentration of ATP was measured as 6.3 μmol/gDW. The intracellular concentrations measured by LC-MSMS were around 1.9, 0.4, 0.5, 0.2, 2.0, 1.4, and 0.07 μmol/gDW for G6P, G1P, F6P, F16P, 2+3PG, PEP, and PYR respectively (table 2). All intracellular concentrations seemed to decrease significantly over the course of about 1000 hours of the fermentation.

**Discussion**

Our goal was the development of a rapid sampling procedure for small samples which allowed sterile sample withdrawal from laboratory fermentors. Ease of handling and accuracy of the results were the major concerns during the development. Since sampling for intracellular measurements requires a subsequent application of sample withdrawal and quenching, sample processing, extraction, drying and metabolite analysis, each step should be verified and possibly optimised. Unfortunately, most steps can not be investigated individually, their precision had to be deduced from the metabolite concentration as measured with either enzymatic analysis or via LC-MSMS. The LC-MSMS method proved to be extremely helpful despite its early development stage, due to its ability to simultaneously determine multiple metabolite concentrations and thus creating large data sets useful for statistical analysis. The results of the statistical error analysis were fed back to set the priorities for the further development of the sampling process.

The simple construction of the sample port prevented clogging and enabled sterile sampling for long periods, both of which were problematic with valve systems containing movable parts. The only requirement on the fermentor was a sub-liquid level port and the ability to stand a modest over-pressure of 0.3 bar. The sample tube adapter allowed various types of tubes to be used, this is especially beneficial if different quenching methods should be employed on the same culture. It facilitated shorter intervals between samples than reported previously and ensured together with the automatically timed valve operation a very consistent sample size. The test tubes do not need to be evacuated prior to the sampling, thus eliminating the need of a septum on the tube and a potentially hazardous needle as used by Theobald and Larsson (Larsson and Törnvist, 1996; Theobald et al., 1997). The mechanism nevertheless prevents spillage and aerosol formation as commonly observed in open setup as described by e.g. Schäfer (Schäfer et al., 1999) or Weuster-Botz (Weuster-Botz, 1999), this allows its application on fermentations of genetically modified organisms (GMOs). The complete procedure with sampling, quenching, washing, extraction, and drying could be performed in the same tube. Functionality, costs, ease of operation and possibility of automation are clear advantages of the presented sampling system over previous systems.
The dead volume in the sampling port represents 5% of the sample amount. This compares very well with the values accomplished with the setup of Theobald (Theobald et al., 1997), where a 200 µl dead volume relates to a 3 ml sample. The possible error introduced by the stagnant liquid volume is thus well below 2%, if the gas holdup in the broth (30%) and a residual glucose concentration of 15 mg/l (van Hoek et al., 1999) is taken into account.

Sample handling required a thorough check of the process temperatures, since it can not be assumed that the methanol alone will completely deactivate all enzymes. Therefore, very low temperatures and fast handling are needed to obtain results reflecting the true metabolic state at the moment of sampling. Temperatures below -20°C are commonly regarded (Weuster-Botz, 1999) sufficient to prevent catalysis. Experiments showed that this temperature was easily exceeded by 10 to 20 degrees if no special precautions had been taken. Cooling of all instruments and substances to -40°C prevented this, but care had to be exercised to prevent freezing of the sample and thus destruction of the cell wall.

Vacuum drying has proven itself to be advantageous in several ways. The most eminent advantage is being able to process several samples simultaneously, with the maximum number of samples being solely dependent on the rotor layout. Thus, the vacuum drying lowers the workload for the rapid sampling procedure, since many samples can be treated simultaneously with little or no supervision required during the procedure. The constantly low temperature during the drying makes the process more repeatable and avoids the final temperature increase unavoidable during film evaporation. The low temperatures during drying might be more beneficial in combination with a different extraction method (e.g. chloroform treatment), currently heat-sensitive components are more strongly affected by the ethanol boiling.

The processing of the sample in the same tubes throughout the whole process removes an additional source of error. Both factors together reduced the relative standard deviation of the whole procedure from 15 to 8% based on MSMS measurements. The enzyme-based intracellular measurements of NAD further confirmed these results. There, the standard deviation decreased from 13 to 7% with the switch from film evaporation to vacuum drying. The latter method thus sharply improves the accuracy of the measurements.

Our average results of the intracellular measurements compare very well with the ones of Theobald et al. (Theobald et al., 1997) (table 2). Nevertheless, significant differences showed up for ATP, NADH, PEP, and PYR. Our low pyruvate levels must be attributed to destruction of the compound during extraction. The same might hold for NADH. The low PEP level determined by Theobald et al. was due to its degradation in the perchloric acid (personal communication M. Reuss). Measured ATP levels were consistently lower than reported by Theobald et al. or Franco et al. (Franco et al., 1984). We found a significant fraction of the overall ATP in the extracellular liquid which might lead to an overestimation of the intracellular concentration, if both fractions are not separated. All other values with the exception of G6P reported by Franco et al. were comparable with our measurement done during earlier stages of the fermentations. Error margins of the different measurements were impossible to compare due to insufficient information.

While MSMS measurements done at different times during the chemostat fermentation varied between 0.9 to 2.6 mM/gDW for G6P and 0.27 to 0.73 mM/gDW for F6P, the ratio
Rapid sampling for intracellular metabolites determination

of F6P to G6P remained constant at 0.27 ±6%, similar to the mass action ratio of 0.29 measured by Benevolensky (Benevolensky et al., 1994). The value is close to the thermodynamic equilibrium value under intracellular conditions of about 0.32 ±0.04 (Noltmann, 1971) due to the fast reaction rate of the isomerase. Similarly, the ratio for PEP/(2PG+3PG) can be calculated as 0.43 to 1.12, based on the equilibria data of 0.12 - 0.2 and 4.0 to 6.7 (Visser et al., 2000) for the reactions from 2PG to 3PG and PEP to 2PG respectively. Our measured values exhibit a fairly constant ratio of 0.75 ±13% while the underlying measurements change by a factor two.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>enzymatic</td>
<td>MSMS</td>
<td>enzymatic</td>
</tr>
<tr>
<td></td>
<td>µmol/gDW stdev</td>
<td>µmol/gDW stdev</td>
<td>µmol/gDW +/- 1)</td>
</tr>
<tr>
<td>ATP</td>
<td>6.3 6%</td>
<td>-</td>
<td>8 3.8%</td>
</tr>
<tr>
<td>NAD</td>
<td>2.4 7%</td>
<td>-</td>
<td>3.2 9.4%</td>
</tr>
<tr>
<td>NADH</td>
<td>0.39 10%</td>
<td>-</td>
<td>2.6 7.7%</td>
</tr>
<tr>
<td>NADP</td>
<td>0.18 10%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NADPH</td>
<td>0.05 15%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>G6P</td>
<td>-</td>
<td>0.9 - 2.6 9.7%</td>
<td>2.2 1.1%</td>
</tr>
<tr>
<td>G1P</td>
<td>-</td>
<td>0.28 - 0.55 7.6%</td>
<td>-</td>
</tr>
<tr>
<td>F6P</td>
<td>-</td>
<td>0.27 - 0.73 7.4%</td>
<td>0.4 2.9%</td>
</tr>
<tr>
<td>F16P</td>
<td>-</td>
<td>&lt;0.12 - 0.4 10.0%</td>
<td>0.26 4.6%</td>
</tr>
<tr>
<td>2+3PG</td>
<td>-</td>
<td>1.2 - 2.6 7.8%</td>
<td>3.2 10)</td>
</tr>
<tr>
<td>PEP</td>
<td>-</td>
<td>0.8 - 1.8 6.2%</td>
<td>0.13 7.6%</td>
</tr>
<tr>
<td>PYR</td>
<td>-</td>
<td>&lt;0.06 - 0.09 18.5%</td>
<td>1.9 2.5%</td>
</tr>
</tbody>
</table>

Table II: comparison of intracellular metabolite concentrations under steady state conditions in S. cerevisiae aerobically grown under glucose limitation in a chemostat. Growth rates were 0.05 hr⁻¹ for the presented data, 0.1 hr⁻¹ for Theobald et al. and about 0.05 hr⁻¹ for Franco et al. The range of measured concentration is due to repeated sampling on the same chemostat over a period of 1000 hours. Fermentation broth was quenched with cold methanol (presented data) followed by subsequent separation of cells and extracellular liquid or with cold perchloric acid (Theobald, Franco).

The measurements of the glycolytic intermediates with the exception of pyruvate was very reliable with about 8% standard deviation, considering the diversity of different steps in the final measuring protocol. Analysis of the error source revealed a good reproducibility of the sampling and sample preparation with a standard deviation of 5.5%. The same value was found for the analysis with HPLC-MSMS. Thus, further improvements will very likely require an optimised analytical protocol.
Conclusions & recommendations

In this article it has been shown that the proposed method of rapid sampling is very well suitable for measurements in a 5-300 seconds time window. Reliable and accurate samples are provided without significant disturbances to the biological state of the fermentation. Large amounts of data as required for the verification of dynamic models or metabolic fingerprinting can be easily supplied by the setup.

Sampling is currently performed largely manually, but both changing of the test tubes and triggering the sampling sequence can be easily automated. The tube adapter is well suited of mechanical insertion of the test tube; the sequential timer can be easily controlled by a computer. The ease of setup, its compatibility with most fermentors, and the flexibility of the current setup make it an ideal system for research with quickly changing requirements.

Sampling processing was greatly simplified and its accuracy improved by the utilisation of vacuum drying and the persistent use of a single tube per sample. The increasing demand for large data sets will ultimately require methods with minimal handling for each sample to decrease the required workload. Quality and work intensity are favourable if compared with other methods.

Acknowledgements

We would like to thank C. Erkelens, R. Kerste, G. Mulder, and J.I. Rhee for their helping hands and technical support, and W. van Gulik for critically reading the manuscript. The work was financed by the European framework IV program.

References


Chapter 8

Summary - Samenvatting
Summary

In the preceding chapters various new experimental and theoretical techniques were devised and applied to quantify the metabolic rates of the yeast *S. cerevisiae* CEN.PK 113-7D grown under various conditions and substrates in carbon-limited aerobic chemostat cultures. The main tool applied for this purpose was the co-factor based metabolic flux analysis, which was briefly introduced and set in the context of other techniques for the quantification of metabolism in the introduction. The scope of the presented work reaches from the initial analysis of the main product of the metabolism, namely biomass, to the quantification of the concentration of some of the glycolytic intermediates of this process.

At the outset, the second chapter of this thesis, the biomass composition of the yeast *S. cerevisiae* was analysed under different growth conditions. Results of various analytical methods for the determination of the biomass’s elemental and molecular constituents were reconciled using statistical analysis and the maximum likelihood method. This resulted in the identification of water as a missing biomass constituent on the list of initial compounds. Subsequent introduction of this additional molecule allowed for the calculation of a consistent and statically acceptable “dry biomass” composition, of which about 4% was tightly bound water. Using this method, the biomass composition was determined for glucose as a carbon and energy substrate over a range of different growth rates. The balanced carbon content was 45.7%, significantly lower than the 48.2% calculated based on the initial measurements of the macromolecules. This clearly indicated the magnitude of possible errors, if the original values would have been used e.g. for the carbon balance of the fermentation. The main constituents of the biomass were protein, carbohydrates, lipids, RNA, phosphate and water with a respective weight based contribution of 40.7, 38.7, 6.1, 6.6, 0.8, and 4.1% of the dry biomass grown at $D = 0.1 \text{ hr}^{-1}$, leading to the unit formula of $\text{CH}_{1.75}\text{N}_{0.146}\text{O}_{0.596}\text{P}_{0.009}\text{S}_{0.0019}$ for glucose grown biomass.

Metabolic flux analysis based on these values was used in the following chapter to predict the aerobic yields of various mutants. A non-compartmented metabolic model was initially used to calculate the energetic parameters of the wild type yeast. Using the maximum yield principle to select between different network options, a $P/O$ ratio of 1.44, a maintenance coefficient $k_{ATP}$ of 0.646 mol ATP/C-mol biomass and a growth related ATP requirement of $m_{ATP} = 55.2$ mmol ATP/C-mol/hr were derived from chemostat cultivations on glucose, ethanol, and acetate at dilution rates from $D = 0.02$ to 0.2 hr$^{-1}$. The metabolic network model with these parameters was then used to forecast for a growth rate of 0.1 hr$^{-1}$ the maximal ethanol uptake of different mutants which were each deleted in one of the enzymes involved in gluconeogenesis: isocitrate lyase (*icl1Δ*), malate synthase (*msΔ*), phosphoenol-pyruvate carboxykinase (*pck1Δ*), and fructose-1,6-bisphosphase (*fbp1Δ*). This allowed the prediction of the biomass concentration in chemostat cultures with varying ratios of ethanol and glucose as mixed carbon substrates. The metabolic model forecasted correctly a sharp decrease, when a certain critical ratio of ethanol to glucose in the medium was exceeded. The maximal ratios, for which the ethanol was completely metabolised, were 47:53, 57:43, and 73:27 on C-mol basis for the *icl1Δ* and *msΔ*, the *pck1Δ*, and the *fbp1Δ* mutants, respectively. Comparison with data gained from chemostat experiments of
the respective mutants reveal a good quantitative fit of the metabolic model prediction with
the measured data for the *icl1Δ*, *mls1Δ*, and *pek1Δ* mutants. Quantitative predictions of
the *fbp1Δ* mutant were less accurate, however, the tendency was forecasted correctly by
the uncompartmented model. This established the first successful quantitative prediction
of mutants based on the maximum yield principle.

To expand the possibilities of the metabolic network analysis, it was found necessary to
include the subcellular compartmentation in the model. The current status of knowledge
on the localisation of the metabolic reactions involved in the central metabolism in *S.
cerevisiae* was given in fourth chapter. Furthermore, the characterised transport proteins
were reviewed and also included to arrive at a model consistent with the existing biological
knowledge. Shuttle systems were used for components for which no transporters were
identified and an existence was deemed doubtful, e.g. the amino group and the C2-carbon
skeleton. The proposed model thus used a isooxovalerate/valine shuttle system for the
mitochondrial import of the amino group. It employed the branched-chain amino acid
transaminase (*BAT1 & BAT2*) and would operate with transporters for isooxovalerate and
valine. Such a system would also readily explain the continuing growth defects of a
*bat1Δbat2Δ* mutant on glucose despite the addition of valine, leucine, and isoleucine.

Special attention was given to the mitochondrial import of the C2-carbon skeleton during
growth on ethanol and acetate. Four possible solutions were discussed. A mitochondrial
acetyl-CoA synthase could not be completely discarded, although some indications pointed
to a cytosolic localisation of both isoenzymes *Acs1p* and *Acs2p*. The same held for the
alternative of an exclusively cytosolic localisation of both enzymes combined with a
cytosolic citrate synthase and a mitochondrial import via an antiport of malate and citrate.
The two other options involved both an acetyl-CoA shuttle system. The operation of a
possible carnitine shuttle could with high probability be excluded for yeast grown on a
synthetic medium. As last option a novel shuttle system was suggest and involved the
mitochondrial and cytosolic isopropyl malate synthases encoded by *LEU4* and *LEU9*,
catalysing the condensation of acetyl-CoA and isoosxovalerate to isopropyl-malate. Based
on these considerations three possible metabolic networks are proposed, each containing
a cytosolic and a mitochondrial compartment, for the aerobic growth of *S. cerevisiae* on
glucose, ethanol, or acetate.

The energetic aspects of compartmentation in *S. cerevisiae* were discussed in chapter five
using the proposed models derived in the previous chapter. Based on chemostat
experiments of yeast grown on glucose and ethanol, the metabolic fluxes and the energetics
of the cell were analysed for growth rates between *D* = 0.02 and 0.2 hr⁻¹. Besides the
metabolic fluxes though the catalytic proteins, this approach allowed a first quantification
of metabolite transport fluxes through the appropriate transport proteins. Based on these
values, it was shown that the use of antiporters are particularly beneficial for growth on C2-
carbons, where their replacements with proton driven active transporters would result in
a decrease of biomass yields of respectively 5 and 8% for ethanol and acetate. The main
cost of mitochondrial transport fluxes were found to be associated with the phosphate
transport and the trafficking of ATP and ADP, which comprised more than 75% of the
required energy for growth on glucose and more than 95% for growth on ethanol and
acetate. Furthermore, the transport costs for phosphate and the adenosine nucleotides led
to different energetic values for mitochondrial and cytosolic ATP. It was shown, that the
energetic costs for the regeneration of approximately ten molecules of cytosolic ATP through the activities of the respiratory chain equals that of thirteen molecules of mitochondrial ATP.

The compartmentation had furthermore implications on the energetic parameters. An existing model with three parameters, as used in the third chapter, was extended by a fourth variable to enable the distinction between the efficiency of the respiratory chain and the $F_{1}F_{0}$-ATPase, previously lumped together in the $P/O$ ratio. However, distinction of the two efficiencies failed with the employed data set. Nevertheless, a new maximum value of 1.38 for the theoretical $P/O$ ratio based on cytosolic ATP was derived with ideal efficiencies of both, the respiratory chain and the $F_{1}F_{0}$-ATPase. Two networks with distinct locations of the acetyl-CoA synthase were used to calculate the energetic parameters. The network with exclusively cytosolic acetyl-CoA synthesis gave a $P/O$ value of 1.30, while the one with a mitochondrial reaction yielded a $P/O$ of 1.18 cytosolic ATP per mole oxygen. Simultaneously, $k_{ATP}$ dropped by more than half compared to results based on an uncompartmented model to 0.39 and 0.17 mol ATP/C-mol biomass for the networks I and II respectively, because now transport costs were taken into account. $m_{ATP}$ remained basically unchanged (49 and 43 mmol ATP/C-mol biomass/hr respectively). These values were used to predict yields on acetate, however, this validation of the suitability of the different proposed models were inconclusive, because both models gave excellent predictions of the acetate yield.

In the sixth chapter, the metabolic flux analysis is broadened by the additional application of the isotopomer balances using $^{13}$C-labelling measured by NMR. Combining the co-factor-based and the $^{13}$C-labelling-based metabolic flux analysis, the growth of the wildtype S. cerevisiae CEN.PK113-7D and its isogenic triose phosphate isomerase deletion ($tpiA$) mutant was analysed. The application of both methods together with a systematic statistical analysis of the model fit allowed for the identification of the most probable metabolic network from a shortlist of possible options. Based on the assumption of identical networks for the wildtype and mutant, the network using a mitochondrial pyruvate bypass and an NADPH coupled cytosolic acetaldehyde dehydrogenase gave the best fit. These results indicated a much more important role for the pyruvate decarboxylase than previously thought. The requirement for a methyglyoxal bypass could not be confirmed, but neither could its complete absence. The modelling results gave the optimal fit of a very small flux of 0.07 mol methyglyoxal per mol glucose consumed. The outcome of this study clearly demonstrated the capacities of a combined approach of co-factor and $^{13}$C-labelling-based MFA with statistical analysis for the qualification and quantification of metabolic networks.

The final, seventh chapter prepared the step from the steady state modelling towards the inclusion of dynamic situations by providing a sampling procedure, which enables the determination of intracellular metabolite concentrations. Here, a new sampling technique was developed for cultures of S. cerevisiae. Based on existing methods, improvements were introduced to improve the reproducibility while decreasing the amount of broth sampled. Statistical analysis was used to determine the drying as the most error-prone step of the original protocol. Subsequent replacement through vacuum drying and a precise protocol especially with respect to the temperatures allowed the withdrawal and immediate quenching to -20°C of 1 ml both in less than one second while achieving an 8% standard
deviation for the complete procedure including analysis. The developed technique thus allows for the reliable provision of a multitude of samples and is thus ideally suited for analysis in combination with HPLC-MSMS to provide the data required of the setup and testing of dynamic models of the intracellular metabolism of yeast.
Samenvatting

In de voorgaande hoofdstukken zijn nieuwe experimentele en theoretische technieken ontwikkeld en toegepast voor de kwantificering van de metabole reactiesnelheden van de gist Saccharomyces cerevisiae CEN.PK 113-7D die is gegroeid in substraat beperkte aerobe chemostaat culturen onder verschillende condities en op verschillende substraten. De kwantificering van deze reactiesnelheden is gebaseerd op metabole flux analyse inclusief co-factor balansen. Tevens is een vergelijking gemaakt met andere flux kwantificerings methoden. Het uitgevoerde onderzoek bestrijkt een breed gebied van kwantificering van de biomassasamenstelling (welk het hoofdproduct is van microbiële metabolisme) tot het bepalen van de concentratie van, vooral glycolytische, intracellulaire metabolieten.

In het tweede hoofdstuk van dit proefschrift is de biomassasamenstelling van Saccharomyces cerevisiae onderzocht onder verschillende groeicondities. De resultaten van verschillende analysemethoden van elementaire en macromoleculaire biomassasamenstelling zijn vereffend met een statistische methode (maximum likelihood). Water blijkt een vergeten biomassa-component te zijn en "droge" biomassa bevat circa 4% sterk gebonden water. Deze nieuwe methode is gebruikt om de biomassasamenstelling te meten voor glucose als C-bron als functie van specifieke groeisnelheid. Het geschatte koolstofgehalte bedroeg 45.7% en is daarmee significant lager dan 48.2% berekend uit alleen de macromoleculaire samenstelling. Hieruit blijkt direct de grootte van de gemaakte fouten indien een dergelijke vereffening niet wordt toegepast, hetgeen direct doorwerkt in bijv. de C-balans over het gehele proces. Biomassa bevat (glucose, \( D = 0.1 \text{ h}^{-1} \)) eiwit (40.7%), koolhydraat (38.7%), vet (6.1%), RNA (6.6%) fosfaat (0.8%) en water (4.1%), overeenkomend met een biomassafomule \( \text{CH}_{1.75}\text{N}_{0.14}\text{O}_{0.59}\text{P}_{0.006}\text{S}_{0.0019} \).

Vervolgens is metabole flux analyse gebruikt om de aerobe biomassa yield van verschillende mutanten te voorspellen. Daartoe zijn eerst de energetische parameters van het wildtype gist geïntegreerd met een niet gecompartimenteerd model en is tevens het principe om aan te geven dat het maximum yield gebruik om uit verschillende netwerkopties te kiezen. Uit chemostaat experimenten, uitgevoerd met glucose, ethanol en acetaat als koolstofbron in een breed gebied van verdunnings-snelheden (\( D = 0.02 \text{ tot } 0.2 \text{ h}^{-1} \)), zijn de energetische parameters bepaald als \( P/O = 1.44, m_{\text{ATP}} = 55.2 \text{ mmol ATP/C-mol X uur en } k_{\text{ATP}} = 0.646 \text{ molATP/C-molX} \). Het metabole model met deze parameters is vervolgens gebruikt om voor een specifieke groeisnelheid van 0.1 h⁻¹ de maximale ethanol opname snelheid te voorspellen van diverse mutanten waarin telkens een van de gluconeogenese enzymen is verwijderd: isocitrate lyase (icl1Δ), malate synthase (mls1Δ), phosphoenolpyruvate carboxykinase (pck1Δ) en fructose 1,6-bisphosphatase (fbp1Δ). Daarmee kan de biomassaconcentratie worden voorspeld in chemostat cultures met glucose/ethanol als mengsubstraat. Het metabole model bleek in staat de scherpe afname in biomassaconcentratie bij te hoge ethanol/glucose ratio in de chemostaat voeding correct te voorspellen. Deze maximale ratio op C-mol basis, waarbij ethanol nog volledig worden gebruikt is 47:53, 57:43 en 73:27 voor de (icl1Δ) en (mls1Δ), (pck1Δ) en (fbp1Δ) mutanten. De kwantitatieve overeenstemming tussen model en experiment was bijzonder goed voor alle mutanten.
behalve de \( \text{fhp1}\Delta \) mutant. Voor deze laatste mutant werd echter de trend door het model goed voorspeld. Hiermee is voor het eerst het maximum yield principe succesvol gebruikt voor het kwantitatief voorspellen van mutanten.

Om het metabool netwerk uit te breiden tot vragen van intracellulair transport is het nodig om compartimentatie op te nemen. Daartoe is in hoofdstuk vier een overzicht gegeven van de huidige kennis betreffende de locatie van metabole reacties uit het centraal metabolisme van \textit{Saccharomyces cerevisiae}. Ook zijn transport eiwitten opgenomen in het model om te komen tot een gecompartmenteerde model. Voor componenten waarvoor geen transporteiwitten zijn geïdentificeerd en waarvan het bestaan onzeker werd geacht zijn shuttlesystemen gebruikt. Voorbeelden zijn de aminogroep en het C\(_2\)-koolstofsketolet. Het voorgestelde model bevatte dus een isoöxalaraat/valine shuttle voor de mitochondriële import van de aminogroep. De shuttle omvatte de branch-arm chain amino acid transferase (\textit{BAT1} & \textit{BAT2}) in samenwerking met transporteiwitten voor isoöxalaraat en valine. Dit zou ook de afwezigheid van groei verklaren op glucose in een \( bat1\Delta bat2\Delta \) mutant, ondanks de toevoeging van de vertakte aminozuren (valine, leucine, isoleucine).


In hoofdstuk vijf zijn de drie mogelijke metabole netwerken gebruikt om de fluxen te berekenen uit experimentele chemostat data van aerobe groei op glucose en ethanol voor \( D = 0.02 - 0.2 \ h^{-1} \). Naast reactie fluxen zijn ook de transport fluxen berekend en de energetische aspecten daarvan. Het blijkt dat het gebruik van mitochondriële antiporters energetisch zeer voordelig is voor de groei op C\(_2\)-substraat. Als deze antiporters vervangen zouden worden door pmf gedreven actief transport dan leidt dit tot biomass yields op ethanol of acetaat die 5 resp. 8% lager zijn. De voornaamste energiekosten van mitochondrieel transport zijn gerelateerd aan transport van fosfaat, ATP en ADP en vertegenwoordigen 75 respectievelijk 95% van alle transportkosten bij groei op glucose resp. ethanol/acetaat. De transportkosten voor fosfaat, ATP en ADP leiden tot verschillende energiekosten voor het regenereren van mitochondriële en cytosolaire ATP. Energetisch zijn 13 mitochondriële ATP equivalent aan 10 cytosolaire ATP.

Door de compartimentering veranderen ook de energetische parameters. Het gecompartmenteerde model is uitgebreid met een vierde parameter om onderscheid te kunnen maken tussen de efficiency van elektron transport fosforlyering en de protein ATPase, die voorheen werden gelumpd in de \( P/O \)-ratio parameter. Op basis van de beschikbare experimentele data konden de aparte efficiencies niet worden onderscheiden. Wel is een nieuwe theoretische maximale waarde van \( P/O = 1.38 \) afgeleid voor de regeneratie van cytosolaire ATP. Uiteindelijk zijn voor twee gecompartmenteerde netwerken energetische
parameters bepaald. Het netwerk waarin acetyl CoA synthese alleen in het cytosol plaatsvindt geeft $P/O = 1.30$, $m_{\text{ATP}} = 49$ mmol ATP/C-molX/hr en $k_{\text{ATP}} = 0.39$ mol ATP/C-molX. In het alternatieve model vindt de Acetyl-CoA synthese plaats in het mitochondrium en is $P/O = 1.18$, $m_{\text{ATP}} = 43$ mmol ATP/C-molX/hr en $k_{\text{ATP}} = 0.17$ mol ATP/C-molX. Omdat in het gecompartmenteerde model de transport energiekosten zijn meegenomen is $k_{\text{ATP}}$ meer dan de helft lager dan in het model zonder compartimentering. Het meenemen van de transport energiekosten veranderde daarentegen de $m_{\text{ATP}}$ nauwelijks. Deze waarden werden gebruikt om de biomassa-opbrengst op acetaat te bepalen. Vergelijking van beide modellen gaf echter geen uitsluitend over welke het beste was, omdat beiden voortreffelijke voorspellingen van de biomassa-opbrengst op acetaat gaven.

In hoofdstuk 6 is de metabole flux analyse uitgebreid door het combineren van cofactor balansen met $^{13}$C-gebaseerde isopomeren balansen. De wildtype CEN.PK113-7D is vergeleken met zijn isogene mutant zonder triosephosphate isomerase ($t\pi\Delta$). Toepassing van beide typen balansen, samen met een statistische analyse van kandidaat netwerken, leidt tot identificatie van de meest waarschijnlijke netwerken. Uitgaande van identieke netwerken gaf het netwerk met daarin een mitochondriële pyruvate by pass en een NADPH cytosolair acetaldehyde dehydrogenase een beste fit. Deze resultaten duiden op een veel belangrijkere rol van pyruvate decarboxylase als tot nu toe gedacht. Een mogelijke methylglyoxylate bypass kan niet worden bevestigd noch verworpen. Een zeer kleine flux (0.07 mol per mol opgenomen glucose) werd berekend. Uit dat onderzoek blijkt zeer duidelijk de kracht van gecombineerde cofactor en isopomeren balansen met een statistische analyse voor metabole netwerk analyse en metabole flux analyse.

Het laatste hoofdstuk zeven legt de basis voor de uitbreiding van steady state modellering naar dynamische modelvorming. Er is een monstername procedure ontwikkeld, waarmee intracellulaire metaboliet concentraties kunnen worden bepaald met als doel een betere nauwkeurigheid en minder monstervorm. Met statistische methoden bleek de droogstap tot een grote fout te leiden. Deze stap is vervangen door vacuum drogen en een gedetailleerd protocol vooral wat betreft de temperatuur tijdens quenchening van 1 ml sample tot -20°C in minder dan 1 seconde. Uiteindelijk is een standaard deviatie van 8% bereikt voor de gehele procedure van sampling, quenching, extractie en analyse. Deze techniek is betrouwbaar en levert veel samples en is dus zeer geschikt in combinatie met LC-MSMS om experimentele data te leveren voor het creëren en valideren van dynamische modellen van intracellulaire metabolisme van gist.
Acknowledgments

Many people have contributed to make this book possible, some of them helped directly in the research and are mentioned in the respective chapters. Here again, my heartfelt thanks to these helping hands and minds. But many other were helpful in an indirect but nevertheless important way. I will try to mention these in the approximate order of appearance.

My initial contact with the Kluyverlaboratory was made during my first visit to Delft together with my wife Viviana Sofronitzki, when by a lucky incident, Karel Luyben had some spare time. During this conversation he encouraged me to apply for a position, and a good half year later René van der Heijden picked me and my suitcase up from the station. He introduced me to the lab, my room, and to statistical reconciliation. Nienke Vriezen explained me the basics rules of biotech-research and of living in The Netherlands, Ko Vinke the basics of fermentation. Hans van Dijken has always been the last resource on yeast and allowed me to skim the yeast-group for valuable information on how to grow *S. cerevisiae*, which I duly did: Thanks to Marijke Luttik & Co. Furthermore, he together with Patricia Osseweijer coordinated our EU-cooperation, which I really appreciated and enjoyed.

I fermented in four laboratories and lived in as many offices. All this would have been impossible without great support throughout the building. If I name all the people who gave a hand, I will ultimately miss one. I have to pick a few, which spend an exceptional amount of their time on my problems and stand each for a larger group around them: Sjaak Lispet, who never lost sight of where thing could be found, Joop Houwers, who could get any computer to communicate with a fermentor, Hans Kemper and later also Marcel van den Broek, whom I bothered with computer problems more than I would have liked myself, Rob Kerste, the quick hand in the lab, Cor Ras, Michael Eman, Gertan van Zuijlen, Carla Schotel, ... Sorry, it’s hard to draw the line, many more deserve to be listed here. But then this book would turn into “Happy stories from the Kluverlab”.

A special thanks goes to everyone in the rapid sampling group, I learned a lot during this time, and even though not all of the zillion goals were accomplished, we did a tremendous job. It is alway difficult to tell, who learned more in end: the guide or the guided. I for sure learned much from Isabel Campos, Jacques van Heel, Wouter Berendsen, and Liang Wu. Three of them are now also on their way towards a PhD - Good luck with it!

In my offices Nienke Vriezen, Petr Vrabel, Yi-Qing Cui, Preben Krabben, Philippe Gaudin, and Liang Wu were always ready for a good discussion and some laughter - thanks for the time together. Without the time spent in the “kippenhok” I probably would have finished much earlier, but I believe it was essential for my emotional well-being: Susanne Rudolph, I owe you much more than just many cups of coffee.

Last, but not least, I’m extremely grateful to Sef Heijnen, who supported me during the whole project. I very much appreciated the many great discussions we had, the tremendous amount of time he spend on my work, and the unfailing confidence he had in me.
The Author

Hans Christian Lange grew up in North-Germany. After an education as electrician, he studied process engineering at the Technical University Hamburg-Harburg with emphasis on environmental and bioprocess technology. Part of these studies were done at the University of Waterloo, Canada, with the support of a scholarship from the DAAD. The final thesis on biofiltration of off-gases was carried out again in Canada, this time at the University of Toronto. From there, he moved to The Netherlands to proceed with studies on the metabolism of yeast, which resulted in the here presented work.