

Development and Application of Experimental and Modeling Tools for *In Vivo* Kinetic Analysis in *S. cerevisiae*

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 donderdag 3 november 2005 te 15:30 uur
door

Liang WU

scheikundige ingenieur
geboren te Hanzhong, Shaanxi Province, P. R. China

Dit proefschrift is goedgekeurd door de promotor:

Prof. dr. J.J. Heijnen

Samenstelling promotiecommissie:

| | |
|---------------------------|---|
| Rector Magnificus | Voorzitter |
| Prof. dr. J.J. Heijnen | Technische Universiteit Delft, promotor |
| Dr. B.M. Bakker | Vrije Universiteit Amsterdam |
| Prof. dr. T.C. Portais | INSA Toulouse, France |
| Prof. dr. J.T. Pronk | Technische Universiteit Delft |
| Prof. dr. M. Reuss | University of Stuttgart, Germany |
| Dr. D. Schipper | DSM |
| Prof. dr. J.P. van Dijken | Technische Universiteit Delft, reservelid |

The studies presented in this thesis were performed at the Department of Biotechnology, Delft University of Technology. The research was financially supported by NWO and DSM.

Cover design by Spatial Phenomenon Atelier

ISBN: 90-9020021-5

Contents

| | |
|---|-----------|
| List of Abbreviations | 1 |
| 1 Introduction | 3 |
| 1.1 Metabolic engineering | 3 |
| 1.2 Experimental tools for metabolic engineering | 4 |
| 1.3 Modeling tools for metabolic engineering | 5 |
| 1.4 Aim and outline of the thesis | 12 |
| 2 Determination of <i>in vivo</i> OUR and CER | 21 |
| 2.1 Introduction | 21 |
| 2.2 Theoretical aspects | 23 |
| 2.3 Materials and methods | 28 |
| 2.4 Results and discussion | 29 |
| 2.5 Conclusion | 35 |
| Appendix A | 38 |
| Appendix B | 38 |
| 3 Quantitative Metabolome Analysis by IDMS | 41 |
| 3.1 Introduction | 41 |
| 3.2 Materials and methods | 44 |
| 3.3 Results and discussion | 45 |
| 4 Lin-Log Kinetics for Control Parameter Estimation | 53 |
| 4.1 Introduction | 54 |
| 4.2 Theory | 55 |
| 4.3 Methods | 60 |
| 4.4 Results and discussion | 62 |
| 4.5 Conclusion | 70 |
| Appendix A | 73 |
| Appendix B | 74 |
| 5 A Conceptual Functional Genomics Strategy Using Solely Metabolome Data | 77 |
| 5.1 Introduction | 77 |
| 5.2 Concept of the approach | 79 |
| 5.3 The <i>in silico</i> model | 81 |

| | | |
|----------|---|------------|
| 5.4 | Methods | 82 |
| 5.5 | Results and discussion | 82 |
| | Appendix | 91 |
| 6 | <i>In Vivo</i> Kinetic Analysis in Prolonged Chemostats | 93 |
| 6.1 | Introduction | 93 |
| 6.2 | Materials and methods | 95 |
| 6.3 | Results and discussion | 96 |
| 6.4 | Conclusion | 108 |
| | Appendix | 112 |
| 7 | <i>In Vivo</i> Kinetic Analysis by Carbon, Electron and ATP Balances | 115 |
| 7.1 | Introduction | 116 |
| 7.2 | Materials and methods | 117 |
| 7.3 | Results and discussion | 121 |
| | Appendix | 141 |
| 8 | Future Directions | 143 |
| | Summary | 149 |
| | Samenvatting | 152 |
| | List of Publications | 155 |
| | Curriculum Vitae | 157 |
| | Dankwoord | 159 |

List of Abbreviations

Throughout this thesis the following abbreviations are used for the biological compounds and enzymes.

Metabolites

| | |
|--------|----------------------------|
| AA | acetaldehyde |
| AcCoA | acetyl coenzyme A |
| Act | acetate |
| Ala | alanine |
| Arg | arginine |
| Asn | asparagine |
| Asp | aspartate |
| BPG | 2,3-bisphosphoglycerate |
| Cit | citrate |
| DHAP | dihydroxyacetonphosphate |
| EtOH | ethanol |
| Fum | fumarate |
| F1,6bP | fructose 1,6-bisphosphate |
| F2,6bP | fructose 2,6-bisphosphate |
| F6P | fructose 6-phosphate |
| GAP | glyceraldehyde 3-phosphate |
| Glc | glucose |
| Glm | glutamate |
| Gln | glutamine |
| Glyc | glycerol |
| G1P | glucose 1-phosphate |
| G3P | glycerol 3-phosphate |
| G6P | glucose 6-phosphate |
| His | histadine |
| iCit | isocitrate |
| Ile | isoleucine |
| Leu | leucine |
| Lys | lysine |
| Mal | malate |
| M6P | mannose 6-phosphate |
| OAA | oxaloacetate |
| Orn | ornithine |
| PEP | phosphoenolpyruvate |
| Phe | phenylalanine |
| Pi | orthophosphate |
| Pro | proline |
| Pyr | pyruvate |

| | |
|-------------|-------------------------|
| Ser | serine |
| Suc | succinate |
| Thr | threonine |
| Tyr | tyrosine |
| T6P | trehalose 6-phosphate |
| Val | valine |
| 2PG | 2-phosphoglycerate |
| 3PG | 3-phosphoglycerate |
| 6PG | 6-phosphogluconate |
| α KG | α -ketoglutarate |

Enzymes

| | |
|-------|--|
| AADH | acetaldehyde dehydrogenase |
| ADH | alcohol dehydrogenase |
| ENO | enolase |
| FBA | fructose-bisphosphate aldolase |
| GAPDH | glyceraldehyde-3-phosphate dehydrogenase |
| GPD | glycerol 3-phosphate dehydrogenase |
| GPP | glycerol-3-phosphatase |
| G6PDH | glucose 6-phosphate dehydrogenase |
| HK | hexokinase |
| PDC | pyruvate decarboxylase |
| PFK | 6-phosphofructo-1-kinase |
| PGI | phosphoglucoisomerase |
| PGK | 3-phosphoglycerate kinase |
| PGM | phosphoglycerate mutase |
| PK | pyruvate kinase |
| TPI | triosephosphate isomerase |

Chapter 1

Introduction

1.1 Metabolic engineering

Industrial biotechnology utilizes biocatalysts for the production of useful chemical substances, such as bulk and fine chemicals, pharmaceuticals, bioplastics, vitamins, food additives and bio-fuels. The main driving forces for the development and implementation of bioprocesses, rather than their classical chemical counterparts, are cost reduction and environmental benefits. The former is due to e.g. the use of cheaper feedstock, higher yields, less conversion steps and reduced energy consumption. The latter mainly involves the use of renewable raw materials, such as cellulose and starch, reduced waste production and reduced CO₂ emission, due to lower energy consumptions.

The applied biocatalysts can be enzymes, microbial and animal cells, as well as plants. In particular, microorganisms (bacteria, yeast and fungi) have been employed in well-controlled fermentation processes for an increasing number of products, ranging from inexpensive bulk chemicals (e.g. ethanol and glutamate) to expensive fine chemicals (e.g. vitamin B₁₂). The competitiveness and hence the viability of these processes depends on the continuous improvement of relevant properties of these microorganisms, which is subject to extensive research efforts.

Traditional strain improvement has relied largely on whole cell mutagenesis followed by screening, which has increased product yields remarkably through the years, for instance for penicillin production by *Penicillium chrysogenum* (Lein, 1983). The drawback of this methodology is the time-consuming screening of a large number of mutants and the accumulation of undesirable genetic changes next to the beneficial ones. Conversely, modern recombinant DNA technology allows precise modifications to be introduced to the genetic makeup of cells, leading to alterations in enzyme amounts and properties, gene regulations and/or biochemical reaction network structure. Such targeted genetic changes, with the aim of enhancing cellular functionalities, has been coined ‘metabolic engineering’ (Bailey, 1991), which has emerged into a multi-disciplinary scientific field in the past decade.

Previously, target selection for genetic intervention has largely been directed by biochemical knowledge of the concerned metabolic pathways, accumulated through the years from e.g. kinetic studies of purified enzymes or investigation of certain tran-

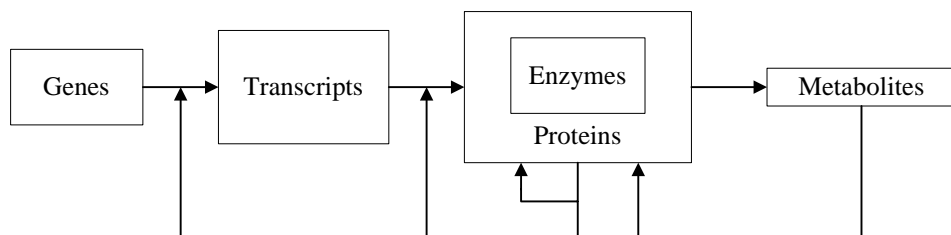


Figure 1.1 Cellular entities and their interactions: essential features of cellular metabolism.

scriptural regulators. This knowledge is often qualitative and incomplete; moreover, it focuses merely on a small part of the entire metabolic and regulatory network and does not sufficiently take into account the interactions between various pathways and cellular entities. It is therefore not surprising that this intuitive approach has met with limited success and has sometimes led to unexpected results. For instance, overexpression of genes for glycolysis and alcohol fermentation failed to increase the rate of ethanol production in *Saccharomyces cerevisiae* (Schaaff et al., 1989).

Hence the so-called metabolic engineering cycle has been proposed (Nielsen, 2001; Stephanopoulos, 1994): the cycle starts with genetic modification (i.e. ‘synthesis’), followed by analysis of the resulting metabolic changes at different cellular levels, which should lead to identification of new targets for a next round of genetic modification (i.e. ‘analysis’). As molecular biology techniques for the ‘synthesis’ step are available for an increasing number of microorganisms, the analysis step becomes crucial to the successful completion of the metabolic engineering cycle. To this end a range of experimental and modeling tools has been developed, which are briefly discussed below.

1.2 Experimental tools for metabolic engineering

The principle features of microbial metabolism are depicted in Fig. 1.1. A number of physical entities can be distinguished, i.e. the genome, transcriptome, proteome and metabolome, which comprises of all the genes, transcripts, proteins and low molecular weight chemical species of the microorganism, respectively. In order for cells to grow and respond to changes in the environment, complex regulations exist, mediated through interactions between and among these entities. For example, gene transcription can be regulated by binding of proteins (e.g. transcriptional factors) to the regulatory elements in the genome; the enzyme capacities can be regulated by post-translational modifications, realized through the action of e.g. protein kinases and phosphatases; in addition, enzyme activities can be regulated by allosteric enzyme-metabolite interactions.

Here the focus will be on experimental methods for the investigation of the individual entities. The high-throughput determination of DNA sequences, gene expressions, protein abundances and metabolite concentrations, respectively, form the basis of a series of experimental ‘omics’-technologies, i.e. genomics, transcriptomics, proteomics and metabolomics.

These experimental tools have evolved to different degrees of maturity in terms

of their analytical spectrum and their practical, large-scale application. Genomics is possibly the best-developed technology, with an ever-increasing number of whole genomes of (micro-) organisms being sequenced. A substantial part of the genes in these sequenced genomes has however yet unknown functions, even for well studied microorganisms such as *S. cerevisiae* (Hughes et al., 2004). Quantification of the expressions of all genes in a genome (relative to a reference), called transcriptome, is performed routinely with DNA microarrays, which are however only commercially available for a limited number of microorganisms of both academic and industrial relevance, such as *S. cerevisiae* and *Escherichia coli*. Compared to the above two ‘genome-wide’ methods, nowadays proteomic and metabolomic technologies cannot yet provide a complete picture for all proteins and metabolites in a cellular system. This is due to a number of factors: for example, proteins are processed and modified in complex ways (e.g. posttranslational modifications), metabolites can be heterogeneous in their chemical properties, and both proteins and metabolites have a dynamic range that spans many orders of magnitude. Nevertheless, considerable advances have been made towards increased through-put and sensitivity by using mass spectrometry based technologies, e.g. isotope-coded affinity tags (ICAT) combined with tandem mass spectrometry (MS/MS) for quantitative proteome analysis (Patterson and Aebersold, 2003) and liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) for intracellular metabolite analysis (Oldiges and Takors, 2005; van Dam et al., 2002).

1.3 Modeling tools for metabolic engineering

The wealth of information obtained from quantitative ‘omics’ techniques however does not directly lead to an understanding of what determines the material flows through different pathways. The material flows in a metabolic system obey the law of mass conservation, which can be described by a set of mass balances.

Consider a metabolic system consisting of m intracellular metabolites with concentrations x_i , m_c extracellular metabolites with concentrations c_i , n reactions with rates v_i , each catalyzed by a different enzyme with a capacity e_i . We thus have:

$$\mathbf{x} = [x_1, x_2, \dots, x_m]^T, \mathbf{c} = [c_1, c_2, \dots, c_{m_c}]^T, \mathbf{e} = [e_1, e_2, \dots, e_n]^T, \mathbf{v} = [v_1, v_2, \dots, v_n]^T$$

The mass balance of intracellular metabolites is:

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v} - \mu c_x \mathbf{i} \quad (1.1)$$

where \mathbf{S} is an $m \times n$ stoichiometry matrix, μ the growth rate, c_x the biomass concentration and \mathbf{i} an $m \times 1$ unit vector. Generally, the metabolic system reaches a steady state, in which the concentrations of intracellular metabolites (\mathbf{x}) are determined by the enzyme capacities (\mathbf{e}), the enzyme kinetic parameters (\mathbf{P}) and the concentrations of extracellular metabolites (\mathbf{c}). The supply and removal of extracellular metabolites can be freely manipulated using proper experimental design, leading to different levels of \mathbf{c} . Hence the intracellular metabolites are also called dependent, while the extracellular ones independent.

The rate of a particular enzyme catalyzed reaction (v_i) is an interplay between the capacity of the enzyme (e_i), the concentrations of the dependent and independent

metabolites (\mathbf{x} , \mathbf{c}) that interact with the enzyme as reactants or allosteric effectors, as well as the interaction strength described by a set of enzyme kinetic parameters (\mathbf{P}). This relationship, usually highly nonlinear, is described by an enzyme kinetic function:

$$v_i = f_i(e_i, \mathbf{x}, \mathbf{c}, \mathbf{P}) \quad (1.2)$$

The enzyme capacity is subject to additional regulatory mechanisms, which can affect the amount by enzyme synthesis and degradation (e.g. through gene expression and translation) and the specific activity by posttranslational processing of the enzyme. The resulting complexity in a large metabolic network clearly necessitates the application of mathematical modeling tools, aiming both at understanding and quantification of *in vivo* interactions (Bailey, 1998). Some of the available tools are briefly discussed below.

Stoichiometric analysis

Stoichiometric analysis is based on mass balances over intracellular metabolites (Eq. 1.1) without employing knowledge of enzyme kinetics Eq. 1.2. Under (quasi) steady state conditions, it holds that:

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v} = \mathbf{0} \quad (1.3)$$

Here the very small term $\mu c_x \mathbf{i}$ is neglected. From Eq. 1.3 intracellular fluxes are estimated using the so-called metabolic flux analysis (Stephanopoulos et al., 1998). With known stoichiometry and a number of measured fluxes, the linear system of equations in Eq. 1.3 can become determined or overdetermined, which allows the direct estimation of intracellular fluxes. When the linear system in Eq. 1.3 is underdetermined, linear programming can be applied with the criteria of, e.g. maximizing growth or metabolite production. In addition to flux estimation for industrial relevant microorganisms (Vallino and Stephanopoulos, 1993; van Gulik et al., 2000), metabolic flux analysis has been applied to estimate energetic parameters (van Gulik et al., 2001), validate metabolic networks (Vanrolleghem et al., 1996), compute metabolic capacities in terms of theoretical yields (van Gulik and Heijnen, 1995), find optimal flux distributions (Varma et al., 1993) and predict phenotypes from genotypes (Edwards et al., 2001; Stuckrath et al., 2002; van Gulik and Heijnen, 1995).

Metabolic flux analysis based on stoichiometry suffers from uncertainty around energy consuming processes and cofactor balances; moreover, fluxes are not observable for parallel pathways and bi-directional reaction steps (Wiechert et al., 2001). These problems can be circumvented by imposing additional constraints, obtained by feeding the metabolic system with labeled precursors (e.g. ^{13}C) and measuring the labeling distribution of intra- and extracellular metabolites by NMR or mass spectrometry. The ^{13}C metabolic flux analysis is a well-developed technique and has been applied to different microbial systems (Gombert et al., 2001; Marx et al., 1996; Sauer et al., 1997; Szyperski, 1995; van Winden et al., 2003).

Kinetic analysis

Stoichiometric analysis provides a description of metabolic fluxes under a particular (quasi-) steady state physiological condition. However, it does not predict how fluxes

will change if the metabolic system is changed by e.g. genetic modifications and therefore does not serve to indicate which genetic modifications should be carried out to change certain fluxes. To this end knowledge of not only the stoichiometric network but also the kinetic interactions are needed, which can be represented by a kinetic model containing a set of ordinary differential equations:

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v}, \quad v_i = f_i(e_i, \mathbf{x}, \mathbf{c}, \mathbf{P}) \quad (1.4)$$

For the set up of a full kinetic model of the metabolic system, two problems need to be addressed: first, a suitable kinetic format (i.e. the function f_i in Eq. 1.4 needs to be chosen and second, the kinetic parameters need to be determined.

Kinetic functions The true, underlying enzyme kinetics can be approximated by two classes of kinetic functions: mechanistic and non-mechanistic. The formulation of mechanistic enzyme kinetic functions requires detailed knowledge of the mechanism of enzyme catalysis. The possibly best-known mechanistic kinetic function is the Michaelis-Menten kinetics (Michaelis and Menten, 1913), in which the reaction rate (v) is a hyperbolic function of the metabolite concentration (x) and proportional to the enzyme capacity (e):

$$v = k_{cat}e \frac{x}{K_m + x}$$

The derivation of mechanistic rate equations for reactions with more than one substrate and product, allosteric and non-allosteric inhibition and activation, as well as cooperative effects is treated in detail in (Cornish-Bowden, 2004). Multiple allosteric control often gives rise to complex mechanistic kinetic functions, such as the phosphofructokinase of *S. cerevisiae*, for which the Monod, Wyman, Changeux model for allosteric enzymes (Monod et al., 1965) was used, including cooperative binding of the substrate F6P, inhibition by ATP and F1,6bP, and allosteric activation by AMP and F2,6bP (Hess and Plesser, 1979; Teusink et al., 2000). In general, mechanistic kinetic functions are believed to have good extrapolation capacity in terms of metabolite concentrations, but are highly nonlinear and contain a large number of parameters, which poses a serious problem for the parameter estimation. For example, for the most common type of reaction occurring in metabolism with two substrates and two products (e.g. $A + B = P + Q$), 10 kinetic parameters are needed to describe the sequential ordered bi-bi mechanism, such as observed for alcohol dehydrogenase (Ganzhorn et al., 1987).

A further general problem is that the use of mechanistic enzyme kinetic functions makes analytical solution of the metabolic network impossible, in terms of steady state metabolite concentrations and fluxes (\mathbf{x} and \mathbf{J} , respectively), when a set of independent metabolite concentrations (\mathbf{c}), enzyme capacities (\mathbf{e}) and kinetic rate expressions (f_i) are given. Instead, the steady state \mathbf{x} and \mathbf{J} can only be obtained through numeric simulation, which limits our understanding of the metabolic network.

Alternatively, non-mechanistic kinetic functions aim to adequately approximate the underlying enzyme kinetics within a limited range of metabolite concentrations, often around a chosen reference state. An overview of the proposed non-mechanistic kinetic functions was given by (Heijnen, 2005). Such approximative kinetic functions should at least be able to reproduce the following essential features of the mechanistic kinetic functions:

- the rate should be proportional to the enzyme capacity
- the kinetic function should be able to describe, to a certain extent, saturation-like behavior, i.e. the reaction rate should show a down concave behavior.

At the same time, the mathematical structure of such approximative kinetic functions should be kept structured and simple for all reactions, should use less kinetic parameters compared to the mechanistic kinetic functions and should preferably allow analytical steady state solutions of the metabolic network.

A number of approximative kinetic functions has been developed over the past years. These include the linear approximations (Heinrich and Rapoport, 1974; Kacser and Burns, 1973), the log-linear kinetics (Hatzimanikatis and Bailey, 1996), the biochemical system theory (Savageau, 1976; Voit, 2000) and kinetic functions containing linear combination of logarithmic metabolite concentrations, i.e. the Mosaic Non-equilibrium Thermodynamics (Westerhoff and van Dam, 1987), the thermokinetic format (Nielsen, 1997), as well as the recently developed lin-log kinetics (Visser and Heijnen, 2002, 2003). It has been concluded that the linear-logarithmic type kinetic functions satisfies the above-mentioned criteria (Heijnen et al., 2004). In general, the linear-logarithmic type kinetic functions can be given by:

$$\mathbf{v} = [\mathbf{e}](\mathbf{a} + \mathbf{p} \ln \mathbf{x} + \mathbf{q} \ln \mathbf{c}) \quad (1.5)$$

Here \mathbf{a} , \mathbf{p} and \mathbf{q} are matrices of the dimension $n \times 1$, $n \times m$ and $n \times m_c$ respectively, which contain kinetic parameters. By introducing a defined reference state (denoted by superscript 0) with fluxes \mathbf{J}^0 , enzyme capacities \mathbf{e}^0 and metabolite concentrations \mathbf{x}^0 and \mathbf{c}^0 , Eq. 1.5 can be rewritten to the reference-based lin-log kinetics (Visser and Heijnen, 2002):

$$\frac{\mathbf{v}}{\mathbf{J}^0} = \left[\frac{\mathbf{e}}{\mathbf{e}^0} \right] \left(\mathbf{i} + \mathbf{E}^{\mathbf{x}^0} \ln \frac{\mathbf{x}}{\mathbf{x}^0} + \mathbf{E}^{\mathbf{c}^0} \ln \frac{\mathbf{c}}{\mathbf{c}^0} \right) \quad (1.6)$$

where the entries in $\mathbf{E}^{\mathbf{x}^0}$ and $\mathbf{E}^{\mathbf{c}^0}$ (ε^{x^0} and ε^{c^0} respectively) are kinetic parameters called elasticities, which are defined in the reference state as scaled local sensitivities of reaction rates v_i towards the metabolite concentrations x_k and c_l :

$$\varepsilon_{ik}^{x^0} = \frac{x_k^0}{v_i^0} \left(\frac{\partial v_i}{\partial x_k} \right)^0, \quad \varepsilon_{il}^{c^0} = \frac{c_l^0}{v_i^0} \left(\frac{\partial v_i}{\partial c_l} \right)^0 \quad (1.7)$$

In a number of *in silico* studies (Visser and Heijnen, 2003; Visser et al., 2004a), the lin-log kinetics has been demonstrated to have adequate approximation capacity in the case of large changes of enzyme and metabolite levels. In addition, it allows analytical steady state solution of the metabolic network. The suitability of the lin-log kinetics to describe experimentally obtained data remains to be demonstrated.

Estimation of kinetic parameters Traditionally, kinetic parameters in mechanistic kinetic functions are estimated through *in vitro* experiments with (partially) purified enzymes or cell extracts. This involves measuring reaction rates while varying metabolite concentrations (usually one at a time) in test tubes, followed by parameter estimation from these data (Cornish-Bowden, 2004). A wealth of kinetic information has been accumulated through the years, which has been systematically stored in e.g. the BRENDA online database.

A caveat of this methodology is that conditions used in test tubes (*in vitro*) deviates considerably from conditions inside the cells (*in vivo*), which likely leads to altered kinetic properties of the studied enzyme. For example, the concentration ratio of enzyme/metabolite in the cells is usually much higher than used in test tubes. *In vivo* effects, such as crowding, channeling and interactions between proteins, are likely to affect enzyme kinetic properties but are not easily examined *in vitro*. A number of models based on *in vitro* derived kinetic parameters failed to reproduce experimental observations (Teusink et al., 2000; Wright and Kelly, 1981).

Hence, there exists the need to obtain kinetic information under *in vivo* conditions. In general, *in vivo* kinetic properties can be exposed through perturbation of the metabolism of intact cells. The perturbation can lead to new (quasi-) steady states, or a short dynamic transient. The first type of perturbation experiments can be called multiple steady state experiments and the second type dynamic experiments.

In multiple steady state experiments, *in vivo* fluxes, enzyme capacities and metabolite concentrations are determined under a sufficient number of different (quasi-) steady states; parameters are obtained by solving Eq. 1.2 with the obtained data. For example, Giersch (1995) established multiple steady states in a reconstituted metabolic system by separately varying enzyme levels and independent metabolite concentrations; the parameters of a linearized kinetic function were subsequently estimated. Multiple steady states can be established *in vivo* by e.g. applying different dilution rates in chemostat cultures (Yang et al., 2003) or changing the capacity of specific enzymes by genetic engineering (Moritz et al., 2000; Petersen et al., 2001). In these *in vivo* studies enzyme kinetic properties were evaluated with fluxes and metabolite concentrations determined *in vivo*, as well as enzyme capacities measured *in vitro*, which requires a considerable experimental effort. Another significant drawback is that the information content might be severely constrained; for example, due to homeostasis the metabolite concentrations might not significantly change between different steady states.

The principle of dynamic experiments is illustrated in Fig. 1.2. This involves perturbation of a metabolic system at (quasi-) steady state by external stimuli followed by measurement of its dynamic metabolic responses. The required experimental in-

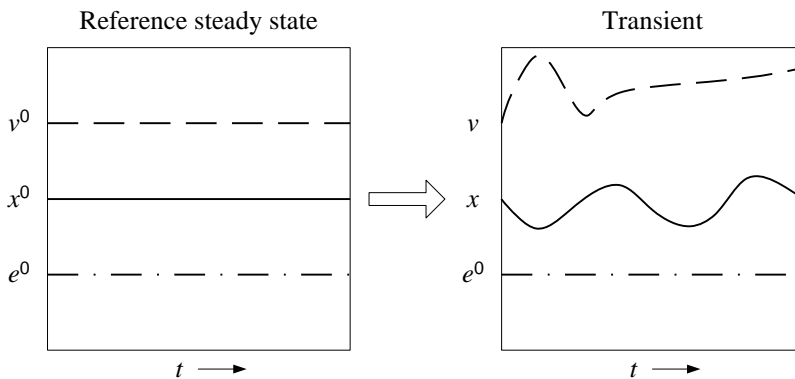


Figure 1.2 Dynamic experiments for the investigation of *in vivo* kinetics.

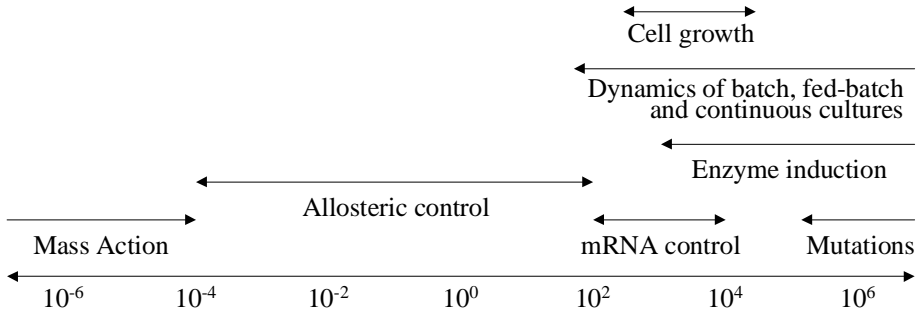


Figure 1.3 Characteristic times of different cellular processes.

formation for *in vivo* kinetic parameter estimation can be significantly reduced by using time scale analysis. Various cellular processes operate on very different time scales, as illustrated in Fig. 1.3. Obviously, by limiting the observation time window to a few hundreds seconds or shorter, it can be assumed that the metabolic responses are primarily triggered by kinetic interactions and that enzyme capacities, which are modified by the slower processes (e.g. enzyme induction), are essentially constant. Thus, contrary to multiple steady state experiments, enzyme capacities do not need to be quantified and measurement of only metabolite concentrations as a function of time allows the estimation of kinetic parameters, from a set of ordinary differential equations:

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v}, \quad v_i = f_i(\mathbf{x}, \mathbf{c}, \mathbf{P})$$

Dynamic experiments have been applied to study the *in vivo* kinetics in the central metabolism (Chassagnole et al., 2002; Theobald et al., 1997; Vaseghi et al., 1999, 2001; Visser et al., 2004b) and product pathways (Oldiges et al., 2004; Schmitz et al., 2002). Parameters of mechanistic kinetic functions have been estimated for the primary metabolism (i.e. glycolysis, TCA cycle and the pentose phosphate pathway) of *S. cerevisiae* (Rizzi et al., 1997; Vaseghi et al., 1999) and *E. coli* (Chassagnole et al., 2002; Degenring et al., 2004).

It is worth mentioning that in the kinetic models obtained so far, the simplifying assumption has often been made that rates of lumped biosynthetic reactions depend on the concentrations of precursors and energy equivalents. Since considerable amounts of biosynthetic precursors (e.g. G6P and pyruvate) are withdrawn from the primary metabolism, it is important, from a mass conservation point of view, to understand the extent of changes of the biosynthetic rates during stimulus response experiments. This is however hampered by yet incomplete measurements of metabolic responses, such as the unquantified dynamic responses of the important carbon sink CO_2 and the redox sink O_2 .

Experimental aspects of dynamic experiments The essential steps of a dynamic experiment are illustrated in Fig. 1.4. The most often used perturbation is a glucose pulse on a carbon-limited culture, which results in a step increase of the residual glucose concentration and triggers large changes in metabolite concentrations and fluxes. In addition, it has been shown that an ethanol pulse on a carbon-limited chemostat of *S. cerevisiae* provides valuable, complementary information

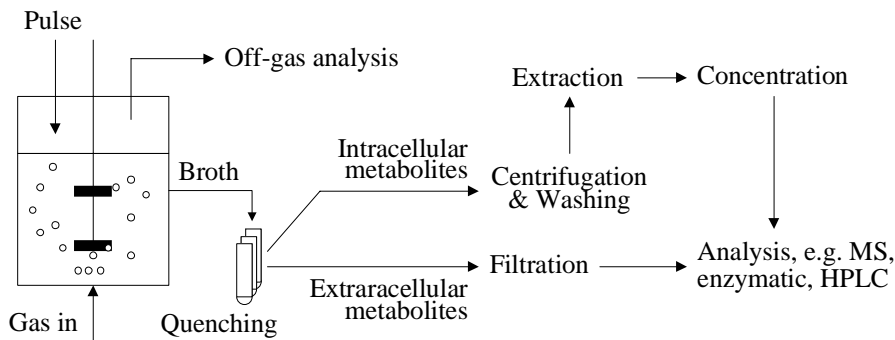


Figure 1.4 Essential steps of a dynamic experiment.

next to glucose pulse experiments (Visser et al., 2004b).

Due to the very short time window of observation (i.e. in the order of seconds to a few minutes), rapid sampling is required to monitor the fast metabolite concentration changes in the metabolic system. Different rapid sampling devices have been developed to realize a sampling frequency of 0.2 – 5 Hz (Buziol et al., 2002; Lange et al., 2001; Schaefer et al., 1999; Theobald et al., 1993; Visser et al., 2002; Weuster-Botz, 1997).

The further turnover of metabolites should be prevented directly after sampling. This is realized by a quenching step, using e.g. methanol/water mixture at -40°C (de Koning and van Dam, 1992) or by immediate cell disruption under extreme pH conditions, using e.g. perchloric acid or KOH (Theobald et al., 1997). When extracellular metabolites are of interest, the fermentation broth is rapidly quenched and filtered (Mashego et al., 2003).

The same metabolite might be present both in cells and in the culture supernatant (secreted by cells or released due to cell death). Since intracellular metabolite concentrations are usually very low (e.g. in the order of $\mu\text{mol/g}_{\text{DW}}$), it is important in this case to separate the biomass from the culture. This is achieved by centrifugation and (repeated) washing under conditions which do not allow metabolite turnover. Cell disruption and metabolite leakage should be minimized during the quenching and washing steps.

Intracellular metabolites are subsequently extracted from the cell pellet after the washing step to obtain a cell-free extract. Frequently applied extraction agents are boiling ethanol (Gonzalez et al., 1997), cold acid/alkaline (Theobald et al., 1997) and cold chloroform (de Koning and van Dam, 1992).

Typically, about a few hundred different intracellular metabolites are present in μM range in the final cell-free extract of a small volume (usually in the milliliter range). Hence, an ideal analytical procedure should combine high specificity, high sensitivity and high throughput with minimal sample usage. In a number of studies, LC-ESI-MS has been shown to be a promising platform for metabolome analysis, which satisfies the mentioned criteria (Buchholz et al., 2001; van Dam et al., 2002).

Several pitfalls can occur in the said experimental and analytical procedures. Cell leakage during quenching and washing, incomplete metabolite extraction, metabolite degradation during sample processing and analysis, as well as ion suppression effects

in LC-ESI-MS (Annesley, 2003), can hamper the unbiased quantification of intracellular metabolite levels. Careful evaluation of these effects, such as standard addition and spiking experiments, which are and unfortunately quite laborious, should be carried out for the different microorganisms of interest (Castrillo et al., 2003; Gonzalez et al., 1997; Hajjaj et al., 1998; Lange et al., 2001; Maharjan and Ferenci, 2003; Wittmann et al., 2004).

Application of kinetic models Once a kinetic model has been obtained, it can be applied for a number of purposes (Wiechert, 2002). First, kinetic models are helpful to organize the available knowledge of enzyme-metabolite interactions. Second, fundamental properties of the metabolic system, e.g. stability, bifurcation or possible oscillations, multiplicity of steady states, can be investigated with kinetic models (Hatzimanikatis and Bailey, 1997; Torres, 1994a; Wolf et al., 2000). In addition, kinetic models allow simulation of the metabolic system, which could assist in understanding of its general properties. When kinetic models are coupled to computational fluid dynamic models, the dynamic metabolic behavior of microorganisms can be studied in industrial scale fermentors, where concentration gradient occurs. This can serve to optimally scale-up fermentation processes (Schmalzriedt et al., 2003).

Furthermore, sensitivity analysis can be performed with a full kinetic model of the metabolic system. Scaled sensitivity coefficients (e.g. elasticities defined in Eq. 1.7) form the basis of the well-developed Metabolic Control Analysis (MCA), which has been widely applied for the quantitative understanding of the distribution of control in metabolic systems (Fell, 1997). These sensitivities, such as the flux and concentration control coefficients, can be straightforwardly computed from a kinetic model (Mauch et al., 1997; Pissara et al., 1996; Torres, 1994b). Alternatively, sensitivities to kinetic parameters can be derived, which could facilitate model reduction (Degenring et al., 2004).

Finally and most importantly, kinetic models are applied for the genetically-based redesign of metabolic systems. A set of optimal enzyme capacities can be computed under the constraint of e.g. maximizing the product flux. This can be combined with additional constraints, such as limited changes in metabolite concentrations and/or total enzyme levels to avoid crowding stress and maintain homeostasis in the cell (Alvarez-Vasquez et al., 2000; Mauch et al., 2001; Visser et al., 2004a). Further optimization has been made possible by incorporating a gene expression model (Schmid et al., 2004), or by co-optimizing the structure and strength of allosteric interactions (Hatzimanikatis et al., 1998).

1.4 Aim and outline of the thesis

An *in vivo* kinetic model of cellular metabolism is a crucial part of the successful completion of the metabolic engineering cycle and has been a focus of research in the bioprocess technology group of the Department of Biotechnology at Delft University of Technology, using among others the primary metabolism of *S. cerevisiae* as a model system. The primary metabolism is the key supplier of precursors and energy and redox cofactors, needed for growth and product formation. It can be expected that enhancement of only the product pathway will eventually lead to limited sup-

ply of precursors and cofactors from the primary metabolism, which will therefore become a further metabolic engineering target. Compared to a product pathway, the primary metabolism contains many more enzymes (about 30 vs. 5 – 10 in a typical product pathway) that are tightly regulated; hence, its kinetic analysis is considerably more complex, which necessitates heavy use of both experimental and modeling tools.

Experimental tools for stimulus response experiments, i.e. rapid sampling and analysis of intracellular metabolites by LC-MS/MS, have been established in the bioprocess technology group (Lange et al., 2001; van Dam et al., 2002). In parallel, a structured yet simple approximative kinetic format, i.e. the lin-log kinetics, has been developed (Visser and Heijnen, 2003). These have laid the foundation of the research described in this thesis, which aims at further development and application of these tools towards a kinetic model of the primary metabolism of *S. cerevisiae*.

Chapters 2 and 3 describe new tools to obtain information from the stimulus response experiments. In **chapter 2**, we develop a mathematical method to reconstruct the *in vivo* oxygen uptake rate (OUR) and carbon dioxide evolution rate (CER) from the off-gas measurements during a dynamic experiment. The OUR and CER convey valuable information of major metabolic processes, such as the respiration, the TCA cycle and the fermentative pathway, but have not been estimated in previous dynamic studies. The proposed method has been improved to incorporate the dissolved oxygen measurement and to take into account variations in the gas flow rate (Bloemen et al., 2003). In **chapter 3**, we develop a novel method for the metabolome quantification by isotope dilution mass spectrometry. This method uses a mixture of fully U-¹³C-labeled metabolites as internal standards in sample processing and LC-ESI-MS/MS analysis, and corrects for metabolite degradation, ion suppression and variations in sample preparation.

Chapters 4 and 5 focus on the application of the lin-log kinetics. **Chapter 4** addresses the problem of estimating elasticities and metabolic control parameters, defined according to the MCA, from multiple steady state data. The traditional MCA-based approach requires infinitesimal changes to be made in the metabolic system, which is experimentally impractical and statistically not sound. It is shown that the application of lin-log kinetics circumvents these problems and leads to a consistent set of elasticities. **Chapter 5** addresses the problem of identifying unknown gene functions from only metabolome data. Previous attempt (i.e. FANCY) compares co-response of metabolites in strains deleted for known and unknown genes (Raamsdonk et al., 2001). In our approach, enzyme capacity changes due to genetic alteration are obtained solely from metabolome data, using a lin-log kinetic model. It is shown *in silico* that enzyme capacity changes calculated from metabolome changes lead to sound hypotheses regarding the unknown gene functions.

Chapters 6 and 7 investigate the *in vivo* kinetics in the primary metabolism of *S. cerevisiae* with the developed experimental and modeling tools. **Chapter 6** employs multiple steady state measurements to identify relevant kinetic mechanisms in prolonged chemostat cultivation (70 – 95 generations) of *S. cerevisiae*, where constant fluxes has been observed accompanied by large changes in enzyme and metabolite levels (Mashego et al., 2005). These measurements are used to estimate elasticities of a lin-log kinetic model. It is shown that due to the limited information content of multiple steady state experiments, only linear constraints between the parameters

can be obtained. This prompts the use of dynamic experiments for *in vivo* kinetic parameter estimation. In **chapter 7**, comprehensive metabolome measurements from a glucose pulse experiment are used to construct cumulative mass, redox and energy balances within a transient of 300 seconds. From these balances, important metabolic processes during the various phases of the transient are identified, which is a prerequisite for the construction of kinetic models.

Finally, **chapter 8** discusses the existing bottlenecks and challenges for *in vivo* kinetic analysis.

References

- Alvarez-Vasquez F, Gonzalez-Alcon C, Torres NV. 2000. Metabolism of citric acid production by *Aspergillus niger*: model definition, steady-state analysis and constrained optimization of citric acid production rate. *Biotechnol Bioeng* 70: 82–108.
- Annesley TM. 2003. Ion suppression in mass spectrometry. *Clin Chem* 49: 1041–4.
- Bailey JE. 1991. Toward a science of metabolic engineering. *Science* 252: 1668–75.
- Bailey JE. 1998. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnol Prog* 14: 8–20.
- Bloemen HJJ, Wu L, van Gulik WM, Heijnen JJ, Verhaegen MHG. 2003. Reconstruction of the O₂ uptake rate and CO₂ evolution rate on a time scale of seconds. *AIChE J* 49: 1895–908.
- Buchholz A, Takors R, Wandrey C. 2001. Quantification of intracellular metabolites in *Escherichia coli* K12 using liquid chromatographic-electrospray ionization tandem mass spectrometric techniques. *Anal Biochem* 295: 129–37.
- Buziol S, Bashir I, Baumeister A, Claassen W, Noisommit-Rizzi N, Mailinger W, Reuss M. 2002. New bioreactor-coupled rapid stopped-flow sampling technique for measurements of metabolite dynamics on a subsecond time scale. *Biotechnol Bioeng* 80: 632–6.
- Castrillo JI, Hayes A, Mohammed S, Gaskell SJ, Oliver SG. 2003. An optimized protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry. *Phytochemistry* 62: 929–37.
- Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M. 2002. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol Bioeng* 79: 53–73.
- Cornish-Bowden A. 2004. *Fundamentals of enzyme kinetics*. Portland Press, London.
- de Koning W, van Dam K. 1992. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Anal Biochem* 204: 118–23.
- Degenring D, Froemel C, Dikta G, Takors R. 2004. Sensitivity analysis for the reduction of complex metabolism models. *J Process Control* 14: 729–45.
- Edwards JS, Ibarra RU, Palsson BO. 2001. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol* 19: 125–30.
- Fell D. 1997. *Understanding the control of metabolism*. Portland press, London.

- Ganzhorn AJ, Green DW, Hershey AD, Gould RM, Plapp BV. 1987. Kinetic characterization of yeast alcohol dehydrogenases. Amino acid residue 294 and substrate specificity. *J Biol Chem* 262: 3754–61.
- Giersch C. 1995. Determining elasticities from multiple measurements of flux rates and metabolite concentrations. application of the multiple modulation method to a reconstituted pathway. *Eur J Biochem* 227: 194–201.
- Gombert AK, Moreira dos Santos M, Christensen B, Nielsen J. 2001. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol* 183: 1441–51.
- Gonzalez B, Francois J, Renaud M. 1997. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast* 13: 1347–55.
- Hajjaj H, Blanc PJ, Goma G, Francois J. 1998. Sampling techniques and comparative extraction procedures for quantitative determination of intra- and extracellular metabolites in filamentous fungi. *FEMS Microbiol Lett* 164: 195–200.
- Hatzimanikatis V, Bailey JE. 1996. MCA has more to say. *J Theor Biol* 182: 233–42.
- Hatzimanikatis V, Bailey JE. 1997. Studies on glycolysis. 1. Multiple steady states in bacterial glycolysis. *Chem Eng Sci* 52: 2579–88.
- Hatzimanikatis V, Emmerling M, Sauer U, Bailey JE. 1998. Application of mathematical tools for metabolic design of microbial ethanol production. *Biotechnol Bioeng* 58: 154–61.
- Heijnen JJ. 2005. Approximative kinetic formats used in metabolic network modeling. *Biotechnol Bioeng* 91: 534–45.
- Heijnen JJ, Van Gulik WM, Shimizu H, Stephanopoulos G. 2004. Metabolic flux control analysis of branch points: an improved approach to obtain flux control coefficients from large perturbation data. *Metab Eng* 6: 391–400.
- Heinrich R, Rapoport TA. 1974. A linear steady-state treatment of enzymatic chains. general properties, control and effector strength. *Eur J Biochem* 42: 89–95.
- Hess B, Plesser T. 1979. Temporal and spatial order in biochemical systems. *Ann N Y Acad Sci* 316: 203–13.
- Hughes TR, Robinson MD, Mitsakakis N, Johnston M. 2004. The promise of functional genomics: completing the encyclopedia of a cell. *Curr Opin Microbiol* 7: 546–54.
- Kacser H, Burns JA. 1973. The control of flux. *Symp Soc Exp Biol* 27: 65–104.
- Lange HC, Eman M, van Zuijlen G, Visser D, van Dam JC, Frank J, de Mattos MJ, Heijnen JJ. 2001. Improved rapid sampling for *in vivo* kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 75: 406–15.
- Lein J. 1983. The panlab strain improvement program in overproduction of microbial metabolites. In: Vanek Z, Hostalek Z (eds.), *Overproduction of microbial metabolites*. Butterworth, London, 105–39.
- Maharjan RP, Ferenci T. 2003. Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*. *Anal Biochem* 313: 145–54.
- Marx A, de Graaf AA, Wiechert W, Eggeling L, Sahm H. 1996. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol Bioeng* 49: 111–29.

- Mashego MR, Jansen ML, Vinke JL, van Gulik WM, Heijnen JJ. 2005. Changes in the metabolome of *Saccharomyces cerevisiae* associated with evolution in aerobic glucose-limited chemostats. *FEMS Yeast Res* 5: 419–30.
- Mashego MR, van Gulik WM, Vinke JL, Heijnen JJ. 2003. Critical evaluation of sampling techniques for residual glucose determination in carbon-limited chemostat culture of *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 83: 395–9.
- Mauch K, Arnold S, Reuss M. 1997. Dynamic sensitivity analysis for metabolic systems. *Chem Eng Sci* 52: 2589–98.
- Mauch K, Buziol S, Schmid JW, Reuss M. 2001. Computer aided design of metabolic networks. In: Chemical process control-6 conference, AIChE Symposium Series. Tucson, Arizona.
- Michaelis L, Menten ML. 1913. Die kinetik der invertin wirkung. *Biochem. Zeitschrift* 49: 333–69.
- Monod J, Wyman J, Changeux JP. 1965. On the nature of allosteric transitions: A plausible model. *J Mol Biol* 12: 88–118.
- Moritz B, Striegel K, De Graaf AA, Sahn H. 2000. Kinetic properties of the glucose-6-phosphate and 6-phosphogluconate dehydrogenases from *Corynebacterium glutamicum* and their application for predicting pentose phosphate pathway flux *in vivo*. *Eur J Biochem* 267: 3442–52.
- Nielsen J. 1997. Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates. *Biochem J* 321: 133–8.
- Nielsen J. 2001. Metabolic engineering. *Appl Microbiol Biotechnol* 55: 263–83.
- Oldiges M, Kunze M, Degenring D, Sprenger GA, Takors R. 2004. Stimulation, monitoring, and analysis of pathway dynamics by metabolic profiling in the aromatic amino acid pathway. *Biotechnol Prog* 20: 1623–33.
- Oldiges M, Takors R. 2005. Applying metabolic profiling techniques for stimulus-response experiments: chances and pitfalls. *Adv Biochem Eng Biotechnol* 92: 173–96.
- Patterson SD, Aebersold RH. 2003. Proteomics: the first decade and beyond. *Nat Genet* 33 Suppl: 311–23.
- Petersen S, Mack C, de Graaf AA, Riedel C, Eikmanns BJ, Sahn H. 2001. Metabolic consequences of altered phosphoenolpyruvate carboxykinase activity in *Corynebacterium glutamicum* reveal anaplerotic regulation mechanisms *in vivo*. *Metab Eng* 3: 344–61.
- Pissara PD, Nielsen J, Bazin MJ. 1996. Pathway kinetics and metabolic control analysis of a high-yielding strain of *Penicillium chrysogenum* during fed batch cultivations. *Biotechnol Bioeng* 51: 168–76.
- Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG. 2001. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* 19: 45–50.
- Rizzi M, Baltés M, Theobald U, Reuss M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol Bioeng* 55: 592–608.
- Sauer U, Hatzimanikatis V, Bailey JE, Hochuli M, Szyperski T, Wuthrich K. 1997. Metabolic fluxes in riboflavin-producing *Bacillus subtilis*. *Nat Biotechnol* 15: 448–52.
- Savageau MA. 1976. Biochemical system analysis. Addison-wesley, Reading (MA).

- Schaaff I, Heinisch J, Zimmermann FK. 1989. Overproduction of glycolytic enzymes in yeast. *Yeast* 5: 285–90.
- Schaefer U, Boos W, Takors R, Weuster-Botz D. 1999. Automated sampling device for monitoring intracellular metabolite dynamics. *Anal Biochem* 270: 88–96.
- Schmalzriedt S, Jenne M, Mauch K, Reuss M. 2003. Integration of physiology and fluid dynamics. *Adv Biochem Eng Biotechnol* 80: 19–68.
- Schmid JW, Mauch K, Reuss M, Gilles ED, Kremling A. 2004. Metabolic design based on a coupled gene expression-metabolic network model of tryptophan production in *Escherichia coli*. *Metab Eng* 6: 364–77.
- Schmitz M, Hirsch E, Bongaerts J, Takors R. 2002. Pulse experiments as a prerequisite for the quantification of *in vivo* enzyme kinetics in aromatic amino acid pathway of *Escherichia coli*. *Biotechnol Prog* 18: 935–41.
- Stephanopoulos G. 1994. Metabolic engineering. *Curr Opin Biotechnol* 5: 196–200.
- Stephanopoulos G, Aristidou AA, Nielsen J. 1998. *Metabolic engineering: principles and methodologies*. Academic Press, San Diego.
- Stuckrath I, Lange HC, Kotter P, van Gulik WM, Entian KD, Heijnen JJ. 2002. Characterization of null mutants of the glyoxylate cycle and gluconeogenic enzymes in *S. cerevisiae* through metabolic network modeling verified by chemostat cultivation. *Biotechnol Bioeng* 77: 61–72.
- Szyperski T. 1995. Biosynthetically directed fractional ^{13}C -labeling of proteinogenic amino acids. an efficient analytical tool to investigate intermediary metabolism. *Eur J Biochem* 232: 433–48.
- Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, Snoep JL. 2000. Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem* 267: 5313–29.
- Theobald U, Mailinger W, Baltés M, Reuss M, Rizzi M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng* 55: 305–16.
- Theobald U, Mailinger W, Reuss M, Rizzi M. 1993. *In vivo* analysis of glucose-induced fast changes in yeast adenine nucleotide pool applying a rapid sampling technique. *Anal Biochem* 214: 31–7.
- Torres NV. 1994a. Modeling approach to control of carbohydrate-metabolism during citric acid accumulation by *Aspergillus niger*. 1. Model definition and stability of the steady state. *Biotechnol Bioeng* 44: 104–11.
- Torres NV. 1994b. Modeling approach to control of carbohydrate-metabolism during citric acid accumulation by *Aspergillus niger*. 2. Sensitivity analysis. *Biotechnol Bioeng* 44: 112–8.
- Vallino JJ, Stephanopoulos G. 1993. Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnol Bioeng* 41: 633–46.
- van Dam JC, Eman M, Frank J, Lange HC, van Dedem GW, Heijnen JJ. 2002. Analysis of glycolytic intermediates in *Saccharomyces cerevisiae* using anion exchange chromatography and electrospray ionization with tandem mass spectrometric detection. *Anal Chim Acta* 460: 209–18.
- van Gulik WM, Antoniewicz MR, deLaat WT, Vinke JL, Heijnen JJ. 2001. Energetics

- of growth and penicillin production in a high-producing strain of *Penicillium chrysogenum*. *Biotechnol Bioeng* 72: 185–93.
- van Gulik WM, de Laat WT, Vinke JL, Heijnen JJ. 2000. Application of metabolic flux analysis for the identification of metabolic bottlenecks in the biosynthesis of penicillin-G. *Biotechnol Bioeng* 68: 602–18.
- van Gulik WM, Heijnen JJ. 1995. A metabolic network stoichiometry analysis of microbial growth and product formation. *Biotechnol Bioeng* 48: 681–98.
- van Winden WA, van Gulik WM, Schipper D, Verheijen PJ, Krabben P, Vinke JL, Heijnen JJ. 2003. Metabolic flux and metabolic network analysis of penicillium chrysogenum using 2D [^{13}C , ^1H] COSY NMR measurements and cumulative bondomer simulation. *Biotechnol Bioeng* 83: 75–92.
- Vanrolleghem PA, de Jong-Gubbels P, van Gulik WM, Pronk JT, van Dijken JP, Heijnen S. 1996. Validation of a metabolic network for *Saccharomyces cerevisiae* using mixed substrate studies. *Biotechnol Prog* 12: 434–48.
- Varma A, Boesch BW, Palsson BO. 1993. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. *Appl Environ Microbiol* 59: 2465–73.
- Vaseghi S, Baumeister A, Rizzi M, Reuss M. 1999. *In vivo* dynamics of the pentose phosphate pathway in *Saccharomyces cerevisiae*. *Metab Eng* 1: 128–40.
- Vaseghi S, Macherhammer F, Zibek S, Reuss M. 2001. Signal transduction dynamics of the protein kinase-A/phosphofruktokinase-2 system in *Saccharomyces cerevisiae*. *Metab Eng* 3: 163–72.
- Visser D, Heijnen JJ. 2002. The mathematics of metabolic control analysis revisited. *Metab Eng* 4: 114–23.
- Visser D, Heijnen JJ. 2003. Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metab Eng* 5: 164–76.
- Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ. 2004a. Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. *Metab Eng* 6: 378–90.
- Visser D, van Zuylen GA, van Dam JC, Eman MR, Proll A, Ras C, Wu L, van Gulik WM, Heijnen JJ. 2004b. Analysis of *in vivo* kinetics of glycolysis in aerobic *Saccharomyces cerevisiae* by application of glucose and ethanol pulses. *Biotechnol Bioeng* 88: 157–67.
- Visser D, van Zuylen GA, van Dam JC, Oudshoorn A, Eman MR, Ras C, van Gulik WM, Frank J, van Dedem GW, Heijnen JJ. 2002. Rapid sampling for analysis of *in vivo* kinetics using the bioscope: a system for continuous-pulse experiments. *Biotechnol Bioeng* 79: 674–81.
- Voit EO. 2000. Computational analysis of biochemical systems: a practical guide for biochemists and molecular biologists. Cambridge University Press, Cambridge.
- Westerhoff HV, van Dam K. 1987. Thermodynamics and control of biological free energy transduction. Elsevier, Amsterdam.
- Weuster-Botz D. 1997. Sampling tube device for monitoring intracellular metabolite dynamics. *Anal Biochem* 246: 225–33.
- Wiechert W. 2002. Modeling and simulation: tools for metabolic engineering. *J Biotechnol* 94: 37–63.
- Wiechert W, Mollney M, Petersen S, de Graaf AA. 2001. A universal framework for ^{13}C metabolic flux analysis. *Metab Eng* 3: 265–83.

- Wittmann C, Kromer JO, Kiefer P, Binz T, Heinzle E. 2004. Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. *Anal Biochem* 327: 135–9.
- Wolf J, Passarge J, Somsen OJ, Snoep JL, Heinrich R, Westerhoff HV. 2000. Transduction of intracellular and intercellular dynamics in yeast glycolytic oscillations. *Biophys J* 78: 1145–53.
- Wright BE, Kelly PJ. 1981. Kinetic models of metabolism in intact cells, tissues, and organisms. *Curr Top Cell Regul* 19: 103–58.
- Yang C, Hua Q, Baba T, Mori H, Shimizu K. 2003. Analysis of *Escherichia coli* anaerobic metabolism and its regulation mechanisms from the metabolic responses to altered dilution rates and phosphoenolpyruvate carboxykinase knockout. *Biotechnol Bioeng* 84: 129–44.

Chapter 2

Determination of *in vivo* OUR and CER

Abstract *In vivo* kinetics of *S. cerevisiae* is studied, in a time window of 150 seconds, by analyzing the response of O₂ and CO₂ in the fermentor off-gas after perturbation of chemostat cultures by metabolite pulses. A new mathematical method is presented for the estimation of the *in vivo* oxygen uptake rate (OUR) and carbon dioxide evolution rate (CER) directly from the off-gas data in such perturbation experiments. The mathematical construction allows effective elimination of delay and distortion in the off-gas measurement signal under highly dynamic conditions. A black box model for the fermentor off-gas system is first obtained by system identification, followed by the construction of an optimal linear filter, based on the identified off-gas model. The method is applied to glucose and ethanol pulses performed on chemostat cultures of *S. cerevisiae*. The estimated OUR is shown to be consistent with the independent dissolved oxygen measurement. The estimated *in vivo* OUR and CER provide valuable insights into the complex dynamic behavior of yeast and are essential for the establishment and validation of *in vivo* kinetic models of primary metabolism.

2.1 Introduction

Targeted alteration of metabolic pathways by recombinant DNA techniques allows the redirection of metabolic fluxes towards desired final products (Lessard, 1996). Prediction of the targets and the outcomes of these alterations is impeded, however, by intrinsic nonlinearity of biological systems, which calls for the application of mathematical models to cellular metabolism (Bailey, 1998). Kinetic models, which combine enzyme kinetics with known stoichiometry of metabolic pathways, have proven useful in describing and understanding the global dynamic behavior of mi-

Published as: Wu L, Lange HC, van Gulik WM, Heijnen JJ. 2003. Determination of *in vivo* oxygen uptake and carbon dioxide evolution rates from off-gas measurements under highly dynamic conditions. *Biotechnol Bioeng* 81: 448-58.

crobes (Gombert and Nielsen, 2000). The importance of using *in vivo* enzyme kinetics in this kind of models, instead of *in vitro* kinetics, has been illustrated by Wright et al. (1992) and Teusink et al. (2000).

In vivo enzyme kinetics can be evaluated by pulse experiments, during which a substrate pulse is administered to a steady state chemostat culture (Theobald et al., 1997). The responses of the culture are observed in a short time window, e.g. 100 – 200 seconds after the pulse, to ensure approximately constant enzyme levels. Analysis of intra- and extracellular metabolite concentrations within this time frame allows the estimation of *in vivo* kinetic parameters (Rizzi et al., 1997; Vaseghi et al., 1999), or the validation and diagnosis of the entire kinetic model (Visser et al., 2000).

In addition to metabolite concentrations, continuous measurements of oxygen and carbon dioxide content in the fermentor off-gas or the fermentation broth can be obtained from online sensors, which facilitate in principle the estimation of the *in vivo* oxygen uptake rate (OUR) and carbon dioxide evolution rate (CER). Both are closely related to the primary metabolism, either through the respiratory chain or reactions in the pyruvate branch point, the TCA cycle and the PP pathway. Despite their relevance to kinetic modeling, estimation of the OUR and CER during transient, in the time scale of 100 – 200 seconds, has not been reported previously in the field of metabolic engineering.

Instead, nonstationary OUR estimation has been addressed in wastewater treatment processes, utilizing dissolved oxygen (DO) measurement. Both instrumental (e.g. respirometry) and software sensors have been applied. A continuous respiratory meter suffers from large calculation intervals (typically 60 seconds) (Spanjers et al., 1994). With software sensors, a constant (oxygen) partial pressure in the gas phase is assumed (Carlsson et al., 1994; Holmberg and Olsson, 1989). Recursive least square techniques or Kalman filtering have been applied to estimate the OUR and the oxygen mass transfer coefficient ($k_L a$) simultaneously, the latter being modeled as a function of the gas flow rate. Lindberg and Carlsson (1996) used a polynomial filtering method to reduce the influence of DO sensor dynamics.

In the case of pulse experiments, the assumption of a constant gas phase O₂ and CO₂ partial pressure no longer holds. Sudden changes in the OUR and CER lead to dynamics in the mass transfer between the liquid and gas phase in the fermentor, and consequently rapid changes in the gas partial pressure. Coupled mass balances for O₂ and CO₂ in the liquid and gas phase are therefore required for the reconstruction of OUR and CER during pulse experiments.

The gas phase O₂ and CO₂ partial pressure is monitored as the volume fractions in the fermentor off-gas, while the dissolved oxygen and carbon dioxide (DO and DCO₂) can be measured closer to the source. The DO and DCO₂ measurements have several drawbacks, however:

- DO or DCO₂ sensors are prone to drift and fouling, which cannot be easily corrected by repeated calibration during fermentations.
- The DCO₂ sensors based on a pH difference measurement are impractical for use in continuous cell cultivation (Pattison et al., 2000), while commercial fiber optic DCO₂ probes do not provide sufficient sensitivity in the concentration range of interest (typically under 5% partial pressure of CO₂). The CER estimation therefore solely depends on off-gas CO₂ measurement.

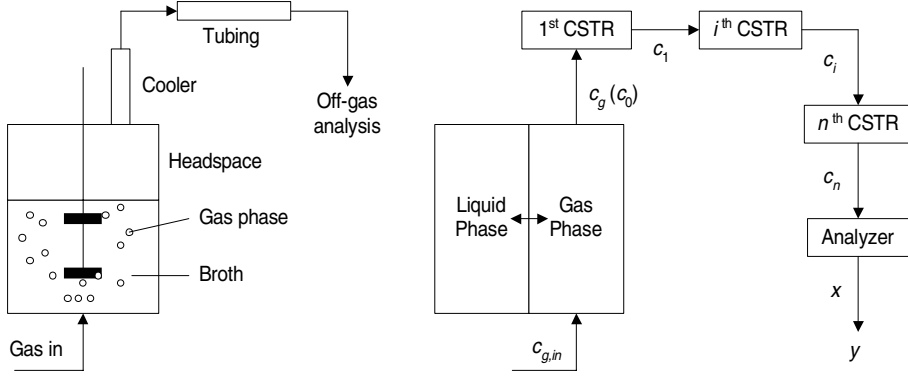


Figure 2.1 Left: schematic depiction of the off-gas system; right: modeling of the off-gas system.

Moreover, off-gas measurements can be easily extended to volatile metabolites by other detection methods, such as mass spectroscopy.

During transient, the off-gas measurement does not reflect the actual gas partial pressures in the fermentor gas phase. The off-gas system in a conventional fermentation setup (see Fig. 2.1) contains large dead volumes, such as headspace and tubing, which cause mixing effects and a time delay. The measured signal is further influenced by sensor dynamics and process and measurement noises.

Taking account of these effects, a new approach is presented in this paper to estimate the OUR and CER during pulse experiments in a time window of 150 seconds, using off-gas measurements. The method is based on the identification of an input-output transfer function model for the off-gas system, followed by the construction of a linear polynomial filter as described by Lindberg and Carlsson (1996). The practical application of this new approach is demonstrated with glucose and ethanol pulse experiments on chemostat cultures of *Saccharomyces cerevisiae*.

2.2 Theoretical aspects

Model development Both the liquid and gas phase (that is the gas present as gas bubbles in the fermentation broth) in the fermentor are described as ideally mixed CSTR's with the following set of differential equations: Gas phase:

$$V_g \frac{dc_g}{dt} = \Phi_g(c_{g,in} - c_g) - k_L a \cdot V_l \left(\frac{c_g}{m} \frac{P}{P_0} - c_l \right) \quad (2.1)$$

Liquid phase:

$$\frac{dc_l}{dt} = \frac{\Phi_l}{V_l} (c_{l,in} - c_l) + k_L a \left(\frac{c_g}{m} \frac{P}{P_0} - c_l \right) + r \quad (2.2)$$

In Eq. 2.1 the gas flow rate in and out of the fermentor is assumed to be constant. At an operating pH of 5, the concentration of carbonate ion is negligible and only dissolved carbon dioxide will contribute to the mass transfer (Royce and Thornhill, 1991). The term $\Phi_l(c_{l,in} - c_l)/V_l$ is neglected in further analysis, since the liquid flow rate is very small.

Further, it is assumed that the complex mixing effects of the off-gas in different compartments (e.g. the headspace, the cooler, the tubing system, etc.) can be approximated by a finite number of CSTR's in a series, all with appropriate dimensions. The model equation for the i^{th} CSTR in the series of CSTR's is given by

$$\frac{dc_i}{dt} = \frac{1}{\tau_i}(c_{i-1} - c_i), \quad i = 1, 2, \dots, N \quad (2.3)$$

where c_i and τ_i are the gas concentration and the residence time of the i^{th} CSTR respectively. c_0 corresponds to the O_2 or CO_2 concentration in the gas bubbles, which enter the headspace. The dynamics of the off-gas analyzer is assumed to be first order and given by

$$\frac{dx}{dt} = \frac{1}{\tau_a}(c_N - x) \quad (2.4)$$

where τ_a is the first order analyzer constant and x the measured gas concentration. The final measurement y is given by

$$y(t) = x(t - k) \quad (2.5)$$

where k is the combined net time delay produced by the dead volumes in the off-gas system. The modeling steps above are shown schematically in Fig. 1.1. Laplace transformation of Eq. 2.1 to 2.5, assuming steady state initial conditions (i.e. all derivatives are zeros in Eq. 2.1 to Eq. 2.5) yields:

$$c'_g(s) = \frac{B_1(s)}{F(s)}c'_{g,in} + \frac{B_2}{F(s)}r'(s) \quad (2.6)$$

$$y'(s) = e^{-sk} \frac{G_n}{G_d(s)}c'_g(s) \quad (2.7)$$

Here, the prime denotes a deviation from the steady state (e.g. $y' = y - y^0$). The definitions of all transfer functions are given in Appendix A. Combining Eq. 2.6 and Eq. 2.7 by eliminating $c'_g(s)$ yields:

$$\begin{aligned} y'(s) &= e^{-sk} \frac{G_n}{G_d(s)} \frac{B_1(s)}{F(s)}c'_{g,in}(s) + e^{-sk} \frac{G_n}{G_d(s)} \frac{B_2}{F(s)}r'(s) \\ &= e^{-sk} \frac{B_c(s)}{A(s)}c'_{g,in}(s) + e^{-sk} \frac{B_r}{A(s)}r'(s) \\ &= G_c(s)c'_{g,in}(s) + G_r(s)r'(s) \end{aligned} \quad (2.8)$$

where

$$\begin{aligned} B_c(s) &= G_n B_1(s), \quad B_r = G_n B_2, \quad A(s) = G_d(s)F(s) \\ G_c(s) &= e^{-sk} \frac{B_c(s)}{A(s)}, \quad G_r(s) = e^{-sk} \frac{B_r}{A(s)} \end{aligned} \quad (2.9)$$

Since off-gas measurements are registered by a digital computer at discrete time intervals, it is convenient to rewrite Eq. 2.8 in a discrete fashion. With the introduction of the forward shift operator q and the backward shift operator q^{-1} , defined by

$$q f(nT) = f(nT + T), \quad q^{-1} f(nT) = f(nT - T) \quad (2.10)$$

the discrete time transfer functions can be obtained by approximation of s in the continuous time transfer functions with Euler's method, when the sampling interval T is short:

$$s \approx \frac{q-1}{T} \quad (2.11)$$

Assuming that corruption of the measurement can be described by white noise $e(t)$, a combination of Eq. 2.8, 2.10 and 2.11 gives the final discrete-time input-output transfer function model of the off-gas system:

$$y'(t) = G_c(q)c'_{g,in}(t) + G_r(q)r'(t) + H(q)e(t) \quad (2.12)$$

where

$$G_c(q) = q^{-k} \frac{B_c(q)}{A(q)}, \quad G_r(q) = q^{-k} \frac{B_r}{A(q)}, \quad H(q) = \frac{C(q)}{D(q)}$$

and A , B , C and D are polynomials in q^{-1} :

$$\begin{aligned} A(q) &= 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} \\ B(q) &= b_1 q^{-1} + \dots + b_{n_b} q^{-n_b} \\ C(q) &= 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c} \\ D(q) &= 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d} \end{aligned}$$

System identification The transfer functions in Eq. 2.12 contain a large number of unknown parameters, e.g. the number of CSTR's in series, their residence time τ_i , etc. (see also Appendix A), the estimation of which will be extremely laborious, if not impossible. We therefore chose to treat the transfer functions as polynomials in q^{-1} with unknown polynomial coefficients, which no longer possess physical relevance. This 'black box' approach allows the estimation of all coefficients with standard system identification techniques (Ljung, 1987). During system identification experiments the off-gas system is excited with known inputs while its response (output) is recorded. A linear quadratic estimate of the polynomial coefficients and time delay can be obtained with the input-output data.

According to Eq. 2.12, the off-gas system has two inputs, namely the gas feed concentration $c_{g,in}$ and the reaction rate r . The choice of the input used for system identification is however limited to $c_{g,in}$, since the reaction rate r cannot be directly and precisely manipulated. Keeping r at a constant level during identification while varying $c_{g,in}$ results in:

$$y'(t) = G_c(q)c'_{g,in}(t) + H(q)e(t) \quad (2.13)$$

which allows the identification of $G_c(q)$ and $H(q)$. $G_r(q)$ can be obtained from the identified $G_c(q)$ as follows: since the two transfer functions share the same denominator $A(q)$ and time delay k , the problem can be reduced to the determination of the scalar B_r , the nominator of $G_r(q)$. The gain \mathcal{G} of the continuous or discrete transfer functions can be determined by taking $s = 0$ or $p = 1$. We thus have:

$$\mathcal{G}_c = 1, \quad \mathcal{G}_r = \tau \quad (2.14)$$

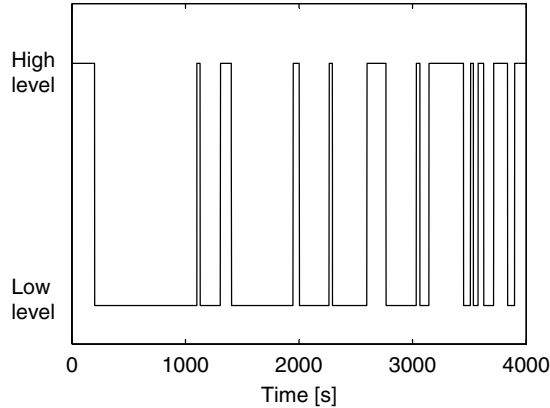


Figure 2.2 A random binary sequence with a fixed low input level and a fixed high input level.

Hence, using Eq. 2.9 and 2.14 we obtain:

$$B_r = \tau \cdot A(s)|_{s=0} \quad \text{or} \quad B_r = \tau \cdot A(q)|_{q=1}$$

The gain relationship allows thus the identification of B_r .

For the actual ‘shape’ of the input, the generalized binary sequence (GBN) has been chosen, illustrated in Fig. 2.2. It is a stochastic signal, which randomly switches between two fixed levels at discrete points in time, with the nonswitching probability p defined as:

$$P(S_t = S_{t-1}) = p, \quad P(S_t \neq S_{t-1}) = 1 - p$$

The expectation of the switching interval T_{sw} is given by

$$E[T_{sw}] = \frac{T_b}{1 - p}$$

where T_b is the length of a fixed basic switching interval.

The GBN can be easily implemented and is at the same time optimal for the statistical information content (Ljung, 1987). Our choice of p was guided by the global guideline outlined by Tulleken (1990), which leads to (sub)optimal GBN design in relation to global dynamic properties of the system to be identified. For a second-order overdamped system, the (sub)optimal nonswitching probability p satisfies

$$\frac{E[T_{sw}]}{\tau_s} = 1 \tag{2.15}$$

where τ_s is the 99% settling time. From the step response of the off-gas system it can be verified that the off-gas system is at least second-order. During identification, the GBN is realized by switching the concentration of O_2 or CO_2 in the gas feed between two constant values. The difference between these two levels has been chosen to cover as much as possible the concentration changes of O_2 or CO_2 in pulse experiments.

The off-gas system during identification should resemble as close as possible the off-gas system during perturbation experiments. This is partially satisfied due

to the time-invariant physical properties of the components of the off-gas system, e.g. the headspace, cooler, tubing system and analyzer. However, care has to be taken that the operational parameters, such as the stirring speed, pH, temperature, pressure, gas flow rate, liquid volume and gas hold-up, are the same during the identification as during the pulse experiments. Additionally, the physical properties (such as density, viscosity, gas holdup and $k_L a$, etc.) of the fermentor liquid phase used for identification should ideally be the same as those of a chemostat broth, which infers the use of a steady state chemostat broth for identification experiments. A serious drawback is that the steady state OUR and CER might be affected by step changes of the input gas concentrations. This is circumvented by using a resting cell culture as an alternative, which possesses approximately the same physical properties as the chemostat broth. The resting cell culture is obtained after switching off the medium feed to the chemostat for several hours, while keeping the other cultivation conditions unchanged.

Optimal linear filtering Based on the identified transfer function $G_r(q)$, a linear filter can be constructed to estimate the actual reaction rates in a least square sense. In the filtering problem the following linear system is defined:

$$y(t) = G(q)u(t) + H(q)e(t) \quad (2.16)$$

where

$$G(q) = q^{-k} \frac{B(q)}{A(q)}, \quad H(q) = \frac{C(q)}{D(q)}, \quad u(t) = \frac{M(q)}{N(q)}v(t)$$

$$Ee(t)^2 = \lambda_e, \quad Ev(t)^2 = \lambda_v, \quad \rho = \lambda_e/\lambda_v$$

Here $e(t)$, $v(t)$ are two independent stationary white and zero-mean noise processes; the noise to signal ratio ρ is defined as the quotient of their variances λ_e and λ_v . q^{-k} corresponds to a pure time delay in the transfer function $G(q)$, so that $B(q)$ does not contain any time delays. With the input model $M(q)/N(q)$, $u(t)$ is modeled as an Auto Regressive Moving Average (ARMA) process.

A polynomial based linear filtering method is used to estimate the input $u(t)$ by minimizing the mean square estimation error (Ahlen and Sternad, 1989):

$$E[\varepsilon(t)^2] = E[u(t) - \hat{u}(t|m)]^2 \quad (2.17)$$

For the estimation of reaction rates, m is negative and the filter corresponds to a fixed lag smoother. The optimal filter can be represented as:

$$\hat{u}(t|m) = \frac{Q(q)}{R(q)}y(t-m) \quad (2.18)$$

The polynomials $Q(q)$ and $R(q)$ that attain the minimum value of Eq. 2.17 are given as:

$$\frac{Q}{R} = \frac{Q_1 DA}{\beta}$$

where the backward shift operator q^{-1} is substituted by the complex argument z^{-1} , which is omitted for convenience.

With the definition of the polynomial P and its conjugate P^* as:

$$\begin{aligned} P &= P(z^{-1}) = 1 + p_1 z^{-1} + \dots + p_{np} z^{-np} \\ P^* &= P(z) = 1 + p_1 z + \dots + p_{np} z^{np} \end{aligned}$$

The polynomial factors Q_1 and β are obtained by solving first a spectral factorization:

$$R\beta\beta_* = MBDM_*B_*D_* + \rho CANC_*A_*N_* \quad (2.19)$$

followed by a Diophantine equation (Ahlen and Sternad, 1989):

$$z^{m+k} M_*B_*D_*MC = R\beta Q_1 + zNL_* \quad (2.20)$$

The choice of noise to signal ratio ρ in Eq. 2.19 is a tradeoff between better tracking of changes in $u(t)$ (when ρ is small) and low variance in estimated $u(t)$ (when ρ is large). The optimal ρ is determined by tuning (Lindberg and Carlsson, 1996). This filtering method is equivalent to stationary Kalman filtering. But compared to the state space formulation of Kalman filtering, the design calculations are simpler, especially for systems with significant time delays and for smoothing problems.

For the off-gas system during perturbation, the reaction rates r are varying while the input gas concentration remains constant, we have according to Eq. 2.12:

$$y'(t) = G_r(q)r'(t) + H(q)e(t)$$

Substitution of $y(t|t-m)$ with $y'(t|t-m)$ in Eq. 2.18 delivers the desired estimate of $r(t)$. To account for the varying nature of $r(t)$, the input model $M(q)/N(q)$ is parameterized as a random walk, which is widely used to model nonstationary time-series (Grewal and Andrews, 1993):

$$\frac{M(q)}{N(q)} = \frac{1}{1-q^{-1}}$$

2.3 Materials and methods

Organism and chemostat cultivation *Saccharomyces cerevisiae* CEN.PK 113-7D was cultivated in carbon limited continuous cultures with a working volume of 4 l in a 7 l fermentor (Applikon, The Netherlands) at a dilution rate of about 0.05 h⁻¹. Conditions applied were a temperature of 30°C, a pH controlled at 5.0, a stirrer speed of 600 rpm, a gas feed (air) flow rate of approximately 3/4 vvm and an overpressure of 0.3 bar. A doubled mineral medium with 27.1 g/l glucose and 1.42 g/l ethanol was used to obtain approximately 15 g biomass (dry weight) per liter in steady state (Lange et al., 2001).

Model identification experiments Identification experiments were performed in a resting cell culture. The (sub)optimal nonswitching probabilities used in model identification experiments were determined by Eq. 2.15, given in Table 2.1.

The GBN was realized by switching between two different levels of O₂ and CO₂ in the gas inlet (Table 2.2). The oxygen and carbon dioxide contents in the off-gas were measured by a NGA200 gas analyzer (Rosemount Analytics, USA). The

Table 2.1 Nonswitching probabilities

| $c_{g,in}$ | T_b (second) | p |
|-----------------|----------------|------|
| O ₂ | 3 | 0.98 |
| CO ₂ | 5 | 0.98 |

Table 2.2 Applied O₂ and CO₂ concentrations in the gas inlet

| | % O ₂ high level | % O ₂ low level | %CO ₂ high level | %CO ₂ low level |
|-----------------|--------------------------------|-------------------------------|--------------------------------|-------------------------------|
| gas/gas mixture | Air | Air+N ₂ | Air+CO ₂ | Air |
| concentration | 20.95 | 18.93 | 1.13 | 0.034 |

dissolved oxygen tension (DOT) was measured by a Mettler Toledo DOT sensor (Mettler-Toledo GmbH, Switzerland). The sampling rate of all measurements was 1 s⁻¹. Estimation of the polynomial coefficients was carried out with the Matlab 5.3 system identification toolbox.

Pulse experiments Under steady state conditions, a pulse of glucose or ethanol was injected into the chemostat with a syringe to give an initial liquid phase concentration of 1 g/l glucose or 0.5 g/l ethanol. The injected volume is negligible compared to the broth volume. Two glucose pulses were conducted on duplicate chemostat cultures G1 and G2. Two ethanol pulses were conducted on duplicate chemostat cultures E1 and E2.

2.4 Results and discussion

Chemostat cultivation Pulse experiments were performed on four separate chemostat cultures. Steady state conditions were checked routinely via biomass concentration and off-gas analysis (Table 2.3). All carbon and redox balances closed within ±3%.

Table 2.3 Summary of steady state chemostat measurements

| Chemostats | D h ⁻¹ | c_x g/l | OUR mmol/g _{DW} /h | CER mmol/g _{DW} /h |
|------------|----------------------|--------------|--------------------------------|--------------------------------|
| G1 | 0.051 | 14.03 | 1.61 | 1.60 |
| G2 | 0.053 | 14.63 | 1.55 | 1.50 |
| E1 | 0.053 | 14.69 | 1.51 | 1.49 |
| E2 | 0.048 | 14.53 | 1.41 | 1.39 |

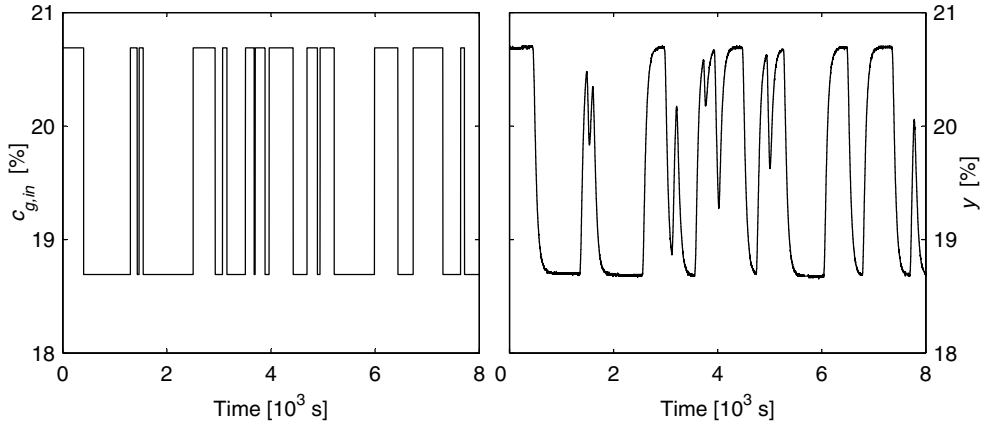


Figure 2.3 Input and output dataset for the identification of a transfer function.

Table 2.4 Polynomial coefficients and their standard deviations of the identified transfer functions

| | O ₂ | | CO ₂ | |
|-------|----------------|-------------------|-----------------|-------------------|
| a_1 | -2.81 | (± 0.004) | -1.96 | (± 0.0002) |
| a_2 | 2.63 | (± 0.009) | 0.96 | (± 0.0002) |
| a_3 | -0.82 | (± 0.004) | | |
| b_1 | 1.44e-4 | ($\pm 3.34e-6$) | 3.35e-4 | ($\pm 1.61e-6$) |
| c_1 | -0.56 | (± 0.01) | -0.23 | (± 0.008) |
| d_1 | -0.98 | (± 0.003) | -0.98 | (± 0.002) |
| d_2 | | | | |
| k | 48 | | 34 | |

System identification A resting cell culture was obtained by switching off the medium feed to the chemostat. This resulted in both the OUR and CER to drop to very low levels (typically $< 5\%$ of the chemostat reaction rates) and no significant change was observed during identification experiments.

A typical identification dataset is shown in Fig. 2.3, from which the distortion of the off-gas measurements is clearly visible. A large number of models with different polynomial orders were evaluated. Model selection was carried out according to the following criteria:

- minimization of mean square prediction error (loss function)
- minimization of parameter confidence intervals;
- fit between the measured output and output simulated by the model;
- whiteness of the prediction errors (residuals);
- independence between the residuals and past inputs;
- a gain close to 1.

The parameters of the selected models and their confidence intervals are sum-

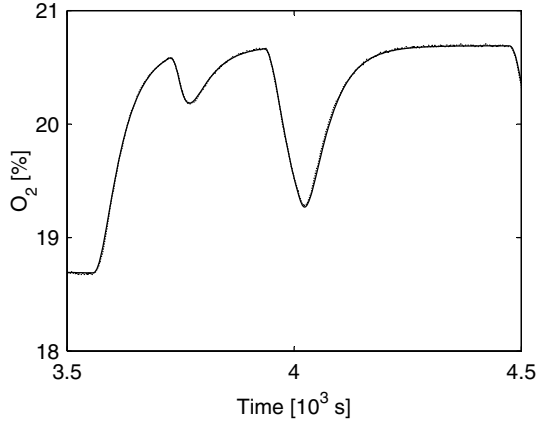


Figure 2.4 Model prediction (—) compared with measured output (···) for an identified transfer function.

marized in Table 2.4. The identified models give good descriptions of the off-gas system, as can be seen from the excellent agreement between the measurement and the simulation (Fig. 2.4).

Optimal linear filtering and verification The pulse responses in a time window of 150 seconds, as measured from the off-gas O_2 and CO_2 content and DO, are plotted in Fig. 2.5 for glucose and ethanol pulses. It is obvious that the off-gas measurements are subject to a large time delay, compared with the fast DO response. The (biomass specific) OUR and CER, estimated from off-gas measurements through the optimal filtering procedure, are plotted in Fig. 2.6. Although the estimated OUR and CER from duplicate cultures are slightly different in magnitude, reproducibility of the trend is satisfactory, which is underlined by the reoccurring patterns.

The estimated OUR and CER can be verified by simulating the identified models with the estimated OUR and CER and compare the outputs, i.e. the simulated off-gas O_2 and CO_2 concentrations, with the measured ones. As can be seen from

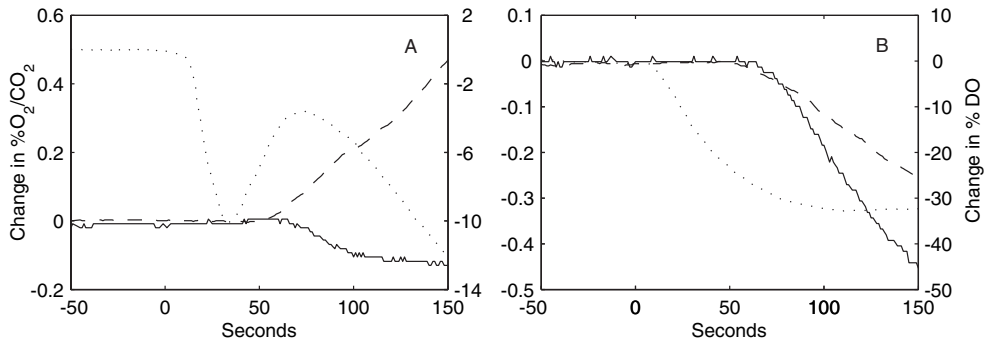


Figure 2.5 Off-gas and DO responses after a glucose pulse on chemostat G1 (A) and an ethanol pulse on chemostat E1 (B). The pulses were both given at $t = 0$. O_2 —; CO_2 - - -; DO, ···.

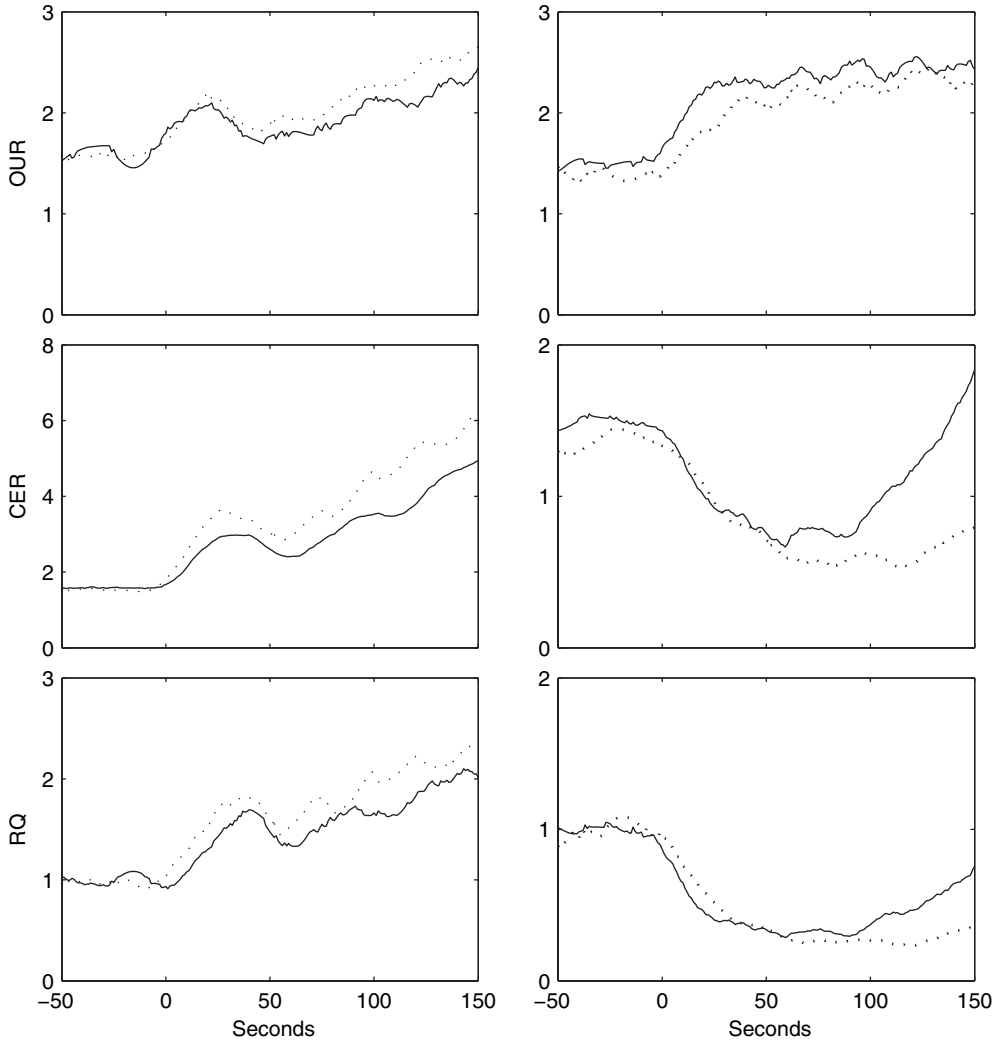


Figure 2.6 Estimated biomass specific OUR and CER in $\text{mmol/g}_{\text{DW}}/\text{h}$ after perturbation of chemostat cultures of *S. cerevisiae*. Left panel: glucose pulses given at $t = 0$ on chemostat G1 (—) and chemostat G2 (\cdots); right panel: ethanol pulses given at $t = 0$ on chemostat E1 (—) and chemostat E2 (\cdots). For OUR, $\rho = 0.01$, $m - k = 20$; for CER, $\rho = 0.001$, $m - k = 20$.

Fig. 2.7, the simulated values fit closely to the measurements.

The initial response of the OUR following the glucose pulses shows a rapid increase followed by a sharp decrease within 50 seconds (see Fig. 2.6). This trend is not obvious in the off-gas O_2 measurement and contradicts what has been expected, i.e. a continuous increase until the capacity of the respiratory chain is saturated. However, the same trend can be observed in the available DO measurement, which supports the estimated initial sharp increase in OUR (as seen from a drop in DO), followed by a decrease in OUR (as seen from the increase in DO). To further confirm

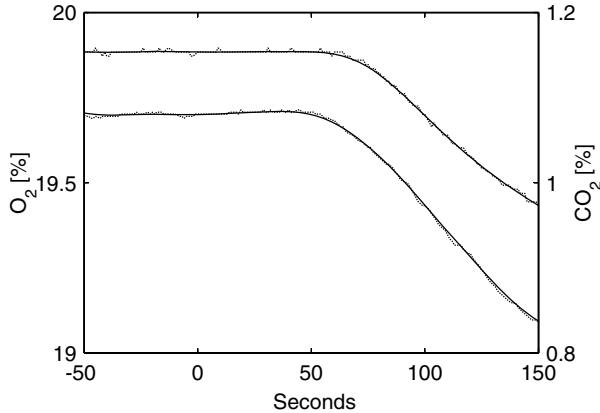


Figure 2.7 Measured (—) and simulated (···) off-gas O_2 and CO_2 concentrations during an ethanol pulse on the chemostat E1. The pulse was given at $t = 0$.

this, we simulated the DO corresponding to the estimated OUR during glucose and ethanol pulses, using the dynamic mass balances Eq. 2.1 and Eq. 2.2. In this calculation, the k_La is assumed to remain constant at the steady state value k_La^0 . As shown in Fig. 2.8, the pattern in the measured DO within 100 seconds after the pulse is structurally well matched by the simulated DO, which validates the unexpected trend in the estimated OUR in the glucose pulses. The same agreement between the measured and simulated DO is shown in Fig. 2.8 for an ethanol pulse. The pattern of estimated OUR is thus consistent with the independent DO measurement.

It should be emphasized that the DO measurement has neither been part of the identified model nor is it used for the OUR calculation. Further, the DO data is not corrected for possible dynamics of the DO probe. This partially explains the deviation of the simulated DO from the measured values in Fig. 2.8. As for the glucose pulses, this deviation amplifies from 100 second onwards. An explanation could be an increase of k_La after the glucose pulse, due to a release of surface-active compounds by the cells. Although we were unable to determine k_La during the highly dynamic period immediately after the pulse, the ‘apparent’ k_La values is calculated from Eq. 2.1 by assuming a pseudo steady state after approximately 7 minutes after the pulse, the result of which suggests a non-constant k_La during glucose pulses (Fig. 2.9).

Because a constant k_La has been assumed in the modeling procedure and embedded in the identified transfer functions, it is important to access the error in the estimated OUR caused by possible changes in k_La during pulse experiments. The error ε can be expressed as:

$$\varepsilon = r_{k_La} - r_{k_La^0}$$

r_{k_La} is the OUR calculated from a true, possibly varying k_La , while $r_{k_La^0}$ denotes the current OUR estimate with a constant k_La assumption. It is shown in Appendix B that ε is within 2% of the current OUR estimate.

The discrepancy between the estimated and simulated DO values and the relatively small error in the estimated OUR can be interpreted as follows: an increase in k_La would, in view of the low solubility of O_2 , immediately lead to a large increase

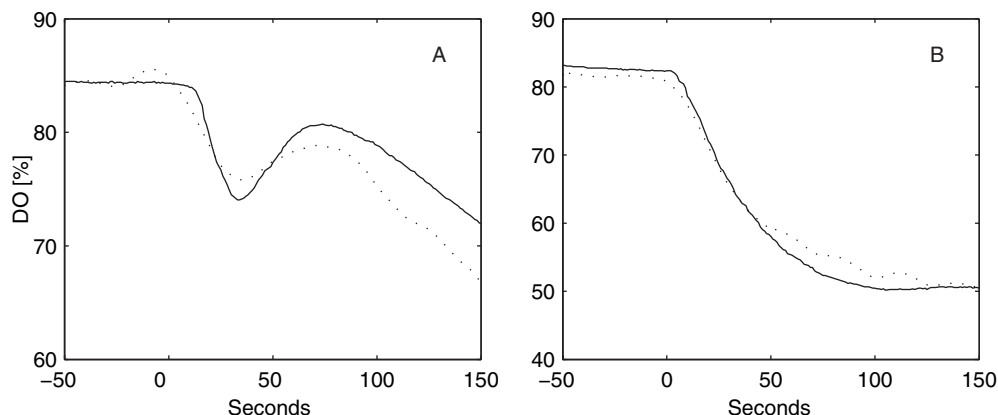


Figure 2.8 Measured (—) and simulated (···) DO during a glucose pulse in chemostat culture G1 (A) and an ethanol pulse in chemostat culture E1 (B), given at $t = 0$. The plotted measured DO is corrected for a delay of 5 seconds.

in the dissolved oxygen concentration, while the O_2 concentration in the gas phase, and hence in the off-gas, is only slightly affected. Thus, the constant $k_L a$ assumption would have resulted in an erroneous OUR estimation if based solely on DO measurement. In the case of CER, the influence of $k_L a$ is even less, due to the high solubility of CO_2 , which strongly reduces the changes in the driving force

$$\frac{c_g}{m} \frac{P}{P_0} - c_l$$

during a pulse. This can be an advantage of estimating OUR and CER using off-gas measurements only, in cases when $k_L a$ is likely to vary and cannot be accurately determined.

Transient OUR and CER responses Two distinct types of response in the direct measurements were triggered by glucose and ethanol as pulse substrate (Fig. 2.5). Both the off-gas O_2 concentration and DO drop shortly after the addition of either glucose or ethanol. This indicates a rapid initial oxygen uptake, which is confirmed by the OUR estimates. The off-gas CO_2 concentration shows however a different trend: it increases during the glucose pulse, but decreases initially during the ethanol pulse.

The complexity of the *in vivo* dynamic behavior during a pulse is better demonstrated by the transient OUR and CER responses. For both glucose pulses (Fig. 2.6), the OUR showed a rapid increase of 50% within 25 seconds, probably due to the direct oxidation of cytosolic NADH produced in the glycolysis. The CER rose about 100%, leading to an RQ close to 2. This indicates ethanol formation, which is confirmed by the measured extracellular ethanol concentration (results not shown). After 25 seconds both OUR and CER decreased, followed by a gradual increase starting from 55 seconds. The reason of this unexpected decrease of both OUR and CER is yet unclear and is subject to further investigation.

For both ethanol pulses (Fig. 2.6), again an initial increase of the OUR was observed, which is of the same magnitude and speed as that observed for the glucose pulse experiments. Contrary to the glucose pulse experiments, the OUR remained

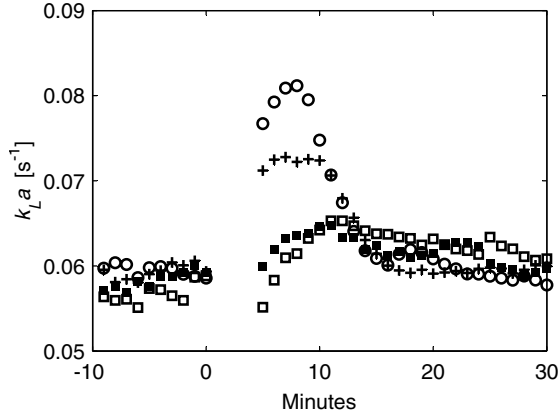


Figure 2.9 Apparent k_{La} based on pseudo steady state assumption. The calculation is left out within 7 minutes after the addition of pulse substrate due to the dynamic nature. G1 (+), G2(o), E1(■), E2(□).

at an elevated level thereafter and increased within 150 seconds towards a plateau of about 2.5 mmol/g_{DW}/h, a value which was also reached in a glucose pulse. However, the CER showed a large unexpected decrease after the ethanol addition and started to increase only after 90 seconds. A possible explanation for the decrease of CER would be as follows: the cytosolic and mitochondrial pyridine nucleotide pool become fully reduced, due to the rapid oxidation of ethanol by alcohol dehydrogenase. This results in a very low level of NAD⁺, which is a substrate of the TCA cycle. A temporary cessation of the TCA cycle would thus lead to the decrease in CER.

2.5 Conclusion

A new method has been developed to estimate the OUR and CER under highly dynamic conditions, taking into account the delay and distortion effects of the off-gas system. The mathematical formulation of the off-gas system, along with the black box simplification of Eq. 2.12, is verified by the capability of the identified models to correctly describe the system response. The linear filtering procedure provides subsequently the OUR and CER estimates, which is consistent with the independent DO measurement and insensitive to k_{La} variations. Good reproducibility of estimated *in vivo* OUR and CER is observed between duplicate cultures.

Distinct types of responses were observed after the addition of either glucose or ethanol as pulse substrate to steady state cultures of *Saccharomyces cerevisiae* grown aerobically on glucose. The dynamics of OUR and CER during glucose and ethanol pulses are resolved for the first time within a time frame of 150 seconds and reveals a complex interplay between various parts of the primary metabolism, which can not yet be adequately accounted for by our current knowledge of its kinetic behavior. The development of reliable *in vivo* kinetic models of microbial metabolism will be strongly benefited by the incorporation of this valuable information, in addition to the measured response of the intra- and extracellular metabolite concentrations.

Nomenclature

| | | |
|--|---|-----------------------|
| $A, B_1, B_c, B_r, C, D, F, G, G_c, G_d, G_r, H, M, N, Q, R$ | transfer functions in s or q | |
| a, b, c, d | polynomial coefficients | |
| B_2, B_r, G_n | scalars in transfer functions | |
| c | concentration of O_2 or CO_2 | mol/m ³ |
| e | zero-mean white noise | |
| \mathcal{G} | gain of transfer function | |
| k | time delay | s |
| $k_L a$ | volumetric mass transfer coefficient | s ⁻¹ |
| m | smoothing lag | s |
| m | equilibrium constant | |
| N | number of CSTR's in series | |
| n_a, n_b, n_c, n_d | order of transfer functions $A(q), B(q), C(q)$ and $D(q)$ | |
| P | pressure in the fermentor | bar |
| P_0 | ambient pressure | bar |
| p | nonswitching probability | |
| q | forward shift operator | |
| r | reaction rate (OUR or CER) | mol/m ³ /s |
| T | sampling interval | s |
| T_b | length of basic switching interval | s |
| T_{sw} | expectation of the switching interval | s |
| u | input | |
| V volume | m ³ | |
| v | zero-mean white noise | |
| x | measured off-gas concentration (undelayed) | mol/m ³ |
| y | measured off-gas concentration (delayed) | mol/m ³ |
| z | complex argument in z transform | |
| ε | estimation error | |
| ε_g | gas hold-up | |
| λ_e, λ_v | variance of white noise e and v | |
| ρ | noise to signal ratio | |
| $\tau_i, \quad i = 1 \dots N$ | residence time in CSTR's in series | s |
| τ_a | first order time constant of off-gas analyzer | s |
| τ_g | gas residence time | s |
| τ_s | 99% settling time | s |
| Φ | flow rate | m ³ /s |
| Subscript | | |
| g | gas phase | |
| l | liquid phase | |
| in | gas or medium feed | |
| Superscript | | |
| ' | change with respect to steady state | |
| 0 | steady state value | |

References

- Ahlen A, Sternad M. 1989. Optimal deconvolution based on polynomial methods. *IEEE trans on acoustic, speech and signal processing* 37: 217–26.
- Bailey JE. 1998. Mathematical modeling and analysis in biochemical engineering: past accomplishments and future opportunities. *Biotechnol Prog* 14: 8–20.
- Carlsson B, Lindberg CF, Hasselblad S, Xu S. 1994. On-line estimation of the respiration rate and the oxygen transfer rate at Kungsängen wastewater treatment plant in Uppsala. *Wat Sci Tech* 30: 255–63.
- Gombert AK, Nielsen J. 2000. Mathematical modelling of metabolism. *Curr Opin Biotechnol* 11: 180–6.
- Grewal MS, Andrews AP. 1993. Kalman filtering: theory and practice. Prentice Hall.
- Holmberg U, Olsson G. 1989. Simultaneous DO control and respiration estimation. *Wat Sci Tech* 21: 1185–95.
- Lange HC, Eman M, van Zuijlen G, Visser D, van Dam JC, Frank J, de Mattos MJ, Heijnen JJ. 2001. Improved rapid sampling for *in vivo* kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 75: 406–15.
- Lessard P. 1996. Metabolic engineering: the concept coalesces. *Nat Biotechnol* 14: 1654–5.
- Lindberg CF, Carlsson B. 1996. Estimation of the respiration rate and oxygen transfer function utilizing a slow DO sensor. *Wat Sci Tech* 33: 325–33.
- Ljung L. 1987. System identification: theory for the user. Prentice Hall PTR, New Jersey.
- Pattison RN, Swamy J, Mendenhall B, Hwang C, Frohlich BT. 2000. Measurement and control of dissolved carbon dioxide in mammalian cell culture processes using an in situ fiber optic chemical sensor. *Biotechnol Prog* 16: 769–74.
- Rizzi M, Baltés M, Theobald U, Reuss M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol Bioeng* 55: 592–608.
- Royce PNC, Thornhill NF. 1991. Estimation of dissolved carbon dioxide concentrations in aerobic fermentations. *AIChE J* 37: 1680–86.
- Spanjers H, Olsson G, Klapwijk A. 1994. Determining short-term biochemical oxygen demand and respiration rate in an aeration tank by using respirometry and estimation. *Wat Res* 28: 1571–83.
- Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, Snoep JL. 2000. Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem* 267: 5313–29.
- Theobald U, Mailinger W, Baltés M, Reuss M, Rizzi M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng* 55: 305–16.
- Tulleken HJAF. 1990. Generalized binary noise test-signal concept for improved identification-experiment design. *Automatica* 26: 37–49.
- Vaseghi S, Baumeister A, Rizzi M, Reuss M. 1999. *In vivo* dynamics of the pentose phosphate pathway in *Saccharomyces cerevisiae*. *Metab Eng* 1: 128–40.
- Visser D, van der Heijden R, Mauch K, Reuss M, Heijnen S. 2000. Tendency modeling: a new approach to obtain simplified kinetic models of metabolism applied to *Saccharomyces*

cerevisiae. Metab Eng 2: 252–75.

Wright BE, Butler MH, Albe KR. 1992. Systems analysis of the tricarboxylic acid cycle in *Dictyostelium discoideum*. I. The basis for model construction. J Biol Chem 267: 3101–5.

Appendix A

$$\begin{aligned}
 B_1(s) &= \frac{1}{\varepsilon_g} \left[s + \frac{k_L a}{\tau_g} \right], & B_2 &= \frac{1}{\varepsilon_g} k_L a \\
 F(s) &= s^2 + \left[k_L a + \frac{1}{\varepsilon_g \tau_g} + \frac{k_L a P_0}{\varepsilon_g m P} \right] s + \frac{k_L a}{\varepsilon_g \tau_g} \\
 G_n &= \frac{1}{\tau_a} \prod_{i=1}^N \frac{1}{\tau_i}, & G_d(s) &= \left(s + \frac{1}{\tau_a} \right) \prod_{i=1}^N \left(s + \frac{1}{\tau_i} \right)
 \end{aligned}$$

The hold-up ε_g is defined as V_g/V_l and τ_g is defined as V_l/F_g .

Appendix B

The estimation error made by constant $k_L a$ assumption is given by:

$$\varepsilon = r_{k_L a} - r_{k_L a^0}$$

Subtracting steady state levels yields:

$$\varepsilon = r'_{k_L a} - r'_{k_L a^0}$$

where the prime denotes a deviation from the steady state.

A useful expression to access the magnitude of ε can be derived as follows: In the noise-free case ($\rho = 0$), the optimal filter reduces to a simple inversion of the transfer function G_r :

$$r'(s) = \frac{1}{G_r(s)} y'(s) \tag{A1}$$

Eq. A1 can be divided into two separate steps: first, c'_g is derived from y' by inverting Eq. 2.7. This step is not dependant on $k_L a$, since the transfer functions in Eq. 2.7 do not contain any $k_L a$ -dependent terms. The reaction rate r' can then be obtained from c'_g by combining mass balances in Eq. 2.1 and Eq. 2.2:

$$r' = \frac{dc'_l}{dt} - [\Phi_g c'_g + V_g \frac{dc'_g}{dt}] \frac{1}{V_l} \tag{A2}$$

Canceling terms containing c'_g in Eq. A2 yields:

$$\varepsilon = r'_{k_L a} - r'_{k_L a^0} = \frac{dc'_{l,k_L a}}{dt} - \frac{dc'_{l,k_L a^0}}{dt}$$

Here $c'_{l,k_L a}$ is the actual liquid phase concentration, subject to a varying $k_L a$, while $c'_{l,k_L a^0}$ denotes the liquid phase concentration if $k_L a$ were constant.

When process and measurement noise is present, estimation of ε is made possible by taking the measured and simulated dissolved oxygen concentration in Fig. 2.9 as approximations for $c_{l,k_L a}$ and $c_{l,k_L a^0}$ respectively. The latter approximation is allowed by realizing that the estimated reaction rate (\hat{r}') corresponds exactly to $r'_{k_L a^0}$. Numerical differentiation of both quantities shows that the error never exceeds 2% of the estimated OUR within 150 seconds after the pulse.

Chapter 3

Quantitative Metabolome Analysis by IDMS

Abstract A novel method was developed for the quantitative analysis of the microbial metabolome, using a mixture of fully U-¹³C-labeled metabolites as internal standards in the metabolite extraction procedure and the subsequent LC-ESI-MS/MS analysis. This mixture of fully U-¹³C labeled metabolites was extracted from biomass of *S. cerevisiae*, cultivated in a fed-batch fermentation on fully U-¹³C-labeled substrates. The obtained labeled cell extract contained in principle the whole yeast metabolome, allowing the quantification of any intracellular metabolite of interest in *S. cerevisiae*. We have applied the labeled cell extract as internal standards in the analysis of glycolytic and TCA cycle intermediates in *S. cerevisiae*, sampled both in steady state and during a transient following a glucose pulse. The use of labeled internal standards effectively reduced errors due to variations occurring in the analysis and sample processing and as a result, the linearity of calibration lines and the precision of measurements were significantly improved. Co-extraction of the labeled cell extract with the samples also eliminates the need of performing elaborate recovery checks for each metabolite to be analyzed. In conclusion, the method presented leads to less workload, more robustness and a higher precision in metabolome analysis.

3.1 Introduction

The integrated approach for the study of microbial physiology requires comprehensive analysis of different functional entities within the cell, such as the levels of mRNA's (transcriptome), proteins (proteome) and low molecular weight metabolites (metabolome) (Delneri et al., 2001). Recently, the importance of metabolomic data has been increasingly recognized in various application fields (Fiehn, 2002), such as the quantification of *in vivo* kinetics using stimulus-response experiments (Buchholz

Published as: Wu L, Mashego MR, van Dam JC, Proell A, Vinke JL, Ras C, van Winden WA, van Gulik WM, Heijnen JJ. 2005. Metabolome Analysis by Isotope Dilution Mass Spectrometry Using Fully U-¹³C-Labeled Cell Extracts as Internal Standards. *Anal Biochem* 336: 164-71.

et al., 2002; Rizzi et al., 1997; Visser, 2002) and qualitative studies in functional genomics (Raamsdonk et al., 2001) and physiology (Tweeddale et al., 1998). Despite the advent of metabolomics, efficient and reliable quantification of (intracellular) metabolite concentrations is still impeded by a number of experimental drawbacks, both in sample preparation and sample analysis.

In practice, sample preparation commences with the fast arrest of metabolism by e.g. quenching in a cold methanol-water solution (de Koning and van Dam, 1992; Gonzalez et al., 1997) followed by an extraction step. Different extraction procedures have been developed, which seek to optimize between complete extraction and limited degradation of metabolites. Established protocols include e.g. boiling ethanol treatment (Gonzalez et al., 1997) and acid/alkaline extraction with freeze-thaw cycles (Theobald et al., 1997). Due to the severe conditions applied (i.e. high temperature or acidity/alkalinity) and instability of a large variety of intracellular species, degradation of metabolites is to be expected and recovery of each metabolite of interest during the sample extraction procedure must be experimentally verified (Gonzalez et al., 1997; Lange et al., 2001). The development of proper sample processing procedures is therefore very time-consuming. Moreover, the overall sample preparation process is elaborate and involves multiple manual handling, which is prone to operator-to-operator and run-to-run variations.

Intracellular metabolite concentrations in the obtained cell extracts can be measured with a range of analytical techniques. The traditional enzymatic and chromatographic methods are being increasingly replaced by high throughput mass spectrometric techniques, which allow the simultaneous identification and quantification of a larger number of metabolites with high selectivity, adequate sensitivity and minimum sample usage (e.g. 5 – 10 μl versus 100 μl or more for enzyme assays). Various studies have shown that liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS) is particularly appropriate for the quantification of phosphorylated carbon compounds, such as encountered in the study of primary metabolic pathways (Buchholz et al., 2001; van Dam et al., 2002). The performance of LC-ESI-MS can however be compromised by ion suppression effects (Annesley, 2003), which renders a different signal response of the analyte due to changes in the sample matrix. The composition of the sample matrix is influenced by e.g. sample preparation procedures and the species of microorganism. The evaluation of ion suppression effects thus encompasses extensive standard addition protocols for each metabolite and each microorganism of interest, which seriously impairs the flexibility of LC-ESI-MS in its implementation and strongly increases the required developmental efforts.

The above-mentioned drawbacks in sample preparation and LC-ESI-MS analysis can be tackled by the isotope dilution technique, where a stable-isotope-analog (i.e. isotopologue, see Hellerstein and Neese, 1999) of the analyte compound is used as an internal standard (IS) in the MS analysis. This technique has found a large variety of applications (Baillie, 1981). Due to the high physical-chemical similarities between the labeled IS and the analyte, degradation during sample preparation, variations in instrumental response and ion suppression effects in LC-ESI-MS can be compensated (Grey et al., 2002; Kuklenyik et al., 2002; Magni et al., 2001; Sojo et al., 2003).

The availability of labeled intracellular metabolites is however scarce and *de novo* synthesis of the isotopologue of interest is often required. Biosynthesis is the preferred

method for the preparation of bio-molecules of complex structure (Matwiyoff and Ott, 1973), e.g. ^{13}C -labeled sugars were obtained by means of photosynthesis from $^{13}\text{CO}_2$ (Kollman et al., 1973). In a number of publications (Evans et al., 2002; Magni et al., 2001), small-scale microbial fermentations were used to yield desired labeled compounds from labeled precursors. For metabolome wide analysis, it is desirable to produce all labeled intracellular metabolites in one stroke. This can be achieved by cultivating organisms exclusively on substrates containing a specific isotope, such as deuterated algae grown on heavy water, in which all the hydrogen present is in the form of the heavy isotope (Katz and Crespi, 1966). Recently, Mashego et al. (2004) have cultivated ^{13}C -labeled biomass by switching the unlabeled feed of a chemostat culture to 100% U- ^{13}C -labeled feed. They found that a short labeling period (4 hours) has led to an incomplete replacement of the unlabeled metabolites in the primary metabolism by U- ^{13}C -labeled ones, possibly due to the relatively slow turnover of unlabeled storage pools. Co-extraction of the partially labeled biomass and unlabeled biomass samples and subsequent LC-ESI-MS/MS analysis provided the ratios of each unlabeled metabolite and its U- ^{13}C -labeled isotopologue (with the necessary correction of the partial labeling in the IS), whereby an improved precision was obtained compared to conventional analysis without IS. However, the adopted methodology still has a number of disadvantages. The production of labeled biomass by chemostat cultivation is rather laborious and much of the labeled material is lost in the effluent from the fermentor. Also the direct use of labeled biomass as IS only provides the relative ratios between unlabeled and U- ^{13}C -labeled metabolites and does not allow determination of absolute metabolite concentrations in samples. The incomplete labeling of the metabolites in the primary metabolism has to be corrected for, increasing the computational complexity and the error in the results. More importantly, certain metabolite pools (especially the conserved moieties such as the adenine and pyridine nucleotides) have a relatively slow turnover rate and will therefore be poorly labeled after a short labeling period, which impedes the isotopic dilution based analysis of these physiologically important metabolites.

We have circumvented the above-mentioned drawbacks by cultivating biomass in a small scale fed-batch fermentation on 100% U- ^{13}C -labeled substrates, from which the labeled metabolites were extracted using proper rapid sampling and quenching procedures. By using a fed-batch culture, all labeled precursors are effectively incorporated into the biomass, resulting in a metabolome-wide labeling that provides in theory fully U- ^{13}C -labeled internal standard for each intracellular metabolite of interest. Moreover, by adding the labeled cell extract to both the unlabeled calibration standards and the samples prior to extraction, absolute metabolite concentrations can be obtained instead of isotopologue ratios. In this paper, we demonstrate the applicability of these labeled cell extracts as IS in LC-ESI-MS/MS analysis and its advantages compared to conventional ^{12}C -based methods for the analysis of glycolytic and TCA cycle intermediates of *S. cerevisiae* in both steady state and transient conditions.

3.2 Materials and methods

Materials U-¹³C₆ Glucose (> 99 at%) and U-¹³C₂ ethanol (> 99 at%) was purchased from Campo Scientific (Veenendaal, The Netherlands). The naturally labeled metabolites G6P, F6P, F1,6bP, G1P, T6P, 2PG, 3PG, Pyr, Cit, αKG, Fum, Suc, Mal were obtained from Sigma (St. Louis, MO, USA). All chemicals used were of analytical grade.

Cultivation of labeled biomass *Saccharomyces cerevisiae* (CEN.PK 113-7D) was cultivated in a 1 l fermentor (Applikon, Schiedam, The Netherlands) equipped with a sampling port at the bottom of the vessel. Throughout the cultivation, the pH was kept at 5, the temperature at 30°C and the overpressure at 0.1 bar. The composition of the mineral medium (0.2 l) used for the batch phase was based on Verduyn et al. (1992) with 14.1 g/l U-¹³C₆ glucose as the sole carbon source. Upon depletion of the carbon source, an aerobic fed-batch phase was started with an exponential feed profile supporting a specific growth rate of 0.05 h⁻¹. The fed-batch medium contained 56.8 g/l U-¹³C₆ glucose, 3.0 g/l U-¹³C₂ ethanol, 16.7 g/l (NH₄)₂SO₄, 10 g/l KH₂PO₄, 1.7 g/l MgSO₄·7H₂O, 3.3 ml/l vitamin solution and 3.3 ml/l solution with trace elements. The vitamin solution contained 0.05 g/l unlabeled D-biotin and 1 g/l unlabeled vitamin B₁. The composition of the trace element solution was according to Verduyn et al. (1992). CO₂ in the air supply was completely removed during the entire cultivation except for the initial batch phase, by passing the air through 2 1-Litre bottles containing 4 M KOH solution connected in series. The stirrer speed and the gas flow rate were adjusted to keep the dissolved oxygen tension above 20%.

Preparation of labeled cell extracts The labeled biomass was sampled three times with a 2-hour-interval between each sampling. To compensate for the volume changes, the feed profile was adjusted after the sampling to ensure a constant growth rate of 0.05 h⁻¹. During each sampling, 100 ml broth was withdrawn from the bottom sampling port of the fermentor into 500 ml of continuously stirred 60 v% methanol kept at -40°C. The 600 ml broth-methanol mixture obtained from each sampling was divided into 100 test tubes in aliquots of 6 ml and kept at -40°C. Centrifugation and washing of the cells in each test tube was carried out as described previously (Mashego et al., 2004). The three sampling runs gave thus 300 test tubes containing cell pellets, which were stored at -80°C no more than 100 hours before being extracted in boiling ethanol (75 v%) as described previously (Mashego et al., 2004). The thus obtained 300 cell extracts were pooled, well mixed and subsequently redistributed in 2 ml vials and stored at -80°C for later use as internal standards.

Preparation of unlabeled biomass samples An aerobic carbon-limited chemostat culture of *S. cerevisiae* (CEN.PK 113-7D) was grown at pH 5, 30°C and a dilution rate of 0.05 h⁻¹ in a 7 l fermentor (Applikon, Schiedam, The Netherlands) with a working volume of 4 l. The feed contained 27.1 g/l glucose and 1.4 g/l ethanol, which supports a biomass concentration of about 15 g_{DW}/l. After about 170 hours, steady state samples were taken within a period of 300 seconds. The chemostat culture was subsequently perturbed by a sudden increase of the extracellular glucose concentration by 1 g/l from the steady state level of 20 mg/l. During the transient

after the perturbation samples were taken within a time window of 300 seconds. The sampling, quenching and extraction procedures followed essentially those described previously (Mashego et al., 2004), except that 100 μl of labeled cell extract was added to each cell pellet prior to the ethanol boiling step. This resulted in a 5 times dilution of the added labeled cell extract in the final sample of 500 μl .

LC-ESI-MS/MS analysis The LC-ESI-MS/MS method for quantification of the glycolytic intermediates in *S. cerevisiae* has been described previously (van Dam et al., 2002). The citric acid cycle compounds were analyzed with the same analysis method and conditions. The metabolites 2PG and 3PG cannot be resolved with the applied analytical procedure, the same held for Cit and iCit. Consequently, only the sum of these compounds (i.e. 2PG+3PG and Cit+iCit) can be determined.

The labeling content of the obtained cell extract was quantified by measuring the concentration of U- ^{13}C -labeled and unlabeled metabolite with a conventional calibration line of unlabeled standards, of which the peak areas corresponding to the mass M+0 were determined (where M is the molecular mass of the metabolite or its fragment). With this procedure we assume that the sensitivity of the analysis is the same for the U- ^{13}C -labeled and ^{12}C -labeled compounds. While we have no pure labeled standards this assumption cannot be checked, but is plausible considering the similarity of physical and chemical behavior of labeled and unlabeled compounds.

For isotope dilution LC-ESI-MS/MS (IDMS) analysis, labeled cell extracts were added to the naturally labeled calibration standards in a 1:4 ratio. The ratio of the unlabeled and U- ^{13}C -labeled metabolites in the samples and the calibration standards was subsequently determined by measuring peak areas corresponding to the masses M+0 and M+N (where N is the number of carbon atoms of the metabolite or its fragment).

3.3 Results and discussion

Production of labeled cell extract With the aim to obtain fully U- ^{13}C -labeled biomass, *S. cerevisiae* (CEN.PK 113-7D) was grown on fully U- ^{13}C -labeled substrates (glucose and ethanol) in an aerobic fed-batch culture and the preceding batch phase. Incorporation of unlabeled carbon was prevented by removing unlabeled CO_2 from the air supply and omitting unlabeled nonessential vitamins from the medium. The only sources of ^{12}C were therefore the inoculum of the batch phase (about 1 ml on a total end volume of 500 ml), the ^{12}C contamination of the labeled substrates (less than 1 at%), the essential vitamins (unlabeled D-biotin and vitamin B₁) present in trace amounts in the medium and the $^{12}\text{CO}_2$ incorporated into the biomass during growth in the initial batch phase, when CO_2 in the air supply was not removed to prevent retardation of growth. Extraction of the thus cultivated and U- ^{13}C labeled biomass yields essentially the whole yeast metabolome and therefore can in principle provide U- ^{13}C -labeled internal standard for each intracellular metabolite of interest.

To assess the extent to which the obtained cell extract was ^{13}C -labeled, the cell extract was analyzed by LC-ESI-MS/MS. The results summarized in Table 3.1 show that the U- ^{13}C enrichments of the metabolites measured are close to 1, indicating a negligible presence of unlabeled fractions, which might interfere with the IDMS analysis.

Table 3.1 Labeling extent of intracellular metabolites in the labeled cell extract^a

| Metabolite | Labeling extent | Metabolite | Labeling extent |
|------------|-----------------|------------|-----------------|
| G6P | 0.99 ± 0.01 | Cit+iCit | 1.00 ± 0.00 |
| F6P | 0.99 ± 0.00 | αKG | 0.99 ± 0.03 |
| F1,6bP | 0.99 ± 0.03 | Suc | 1.00 ± 0.00 |
| 2PG+3PG | 1.00 ± 0.03 | Fum | 1.00 ± 0.00 |
| PEP | 0.99 ± 0.00 | Mal | 1.00 ± 0.00 |
| Pyr | 0.97 ± 0.00 | | |

^a Labeling extent is defined as the concentration ratio of U-¹³C-labeled and the sum of U-¹³C- and ¹²C-labeled metabolites. Concentration ratios and 95% confidence intervals were determined from quadruple measurements of a sample of the labeled cell extract.

Table 3.2 Coefficients of determination (R^2) obtained from IS- and ¹²C-based calibration lines^a

| Metabolite | IS-based cal. line | ¹² C-based cal. line | Metabolite | IS-based cal. line | ¹² C-based cal. line |
|------------|--------------------|---------------------------------|------------|--------------------|---------------------------------|
| G6P | 0.9972 | 0.9928 | Cit+iCit | 0.9927 | 0.9703 |
| F6P | 0.9896 | 0.9724 | αKG | 0.9937 | 0.9903 |
| F1,6bP | 0.9986 | 0.9927 | Suc | 0.9978 | 0.9385 |
| 2PG+3PG | 0.9919 | 0.9695 | Fum | 0.9933 | 0.9700 |
| PEP | 0.9963 | 0.9340 | Mal | 0.9900 | 0.9641 |
| Pyr | 0.9755 | 0.8592 | | | |

^a For a proper comparison linear calibration lines were used without forced zero intercept for the same set of ¹²C-standard concentrations.

Calibration with labeled cell extract as IS Labeled cell extracts were added as internal standard in equal amounts to the unlabeled calibration standards and the peak areas of the U-¹³C and ¹²C metabolites were measured by LC-ESI-MS/MS. Linear calibration lines were obtained when the area ratios between the U-¹³C and ¹²C metabolites are plotted against the known concentrations of ¹²C metabolites in the calibration standard. Since the peak areas of unlabeled metabolites were measured as well in the IDMS analysis, it was also possible to obtain a conventional ¹²C calibration line. The calculated coefficients of determination (R^2) given in Table 3.2 show an improved linearity of the IS-based calibration lines compared to ¹²C-based calibration lines for all metabolites analyzed. In Fig. 3.1 several parity plots are given where the known concentrations in the standards are plotted against the calculated concentrations according to the two different calibration lines. It can be seen in Fig. 3.1 that the outliers (e.g. Suc and Cit+iCit) and nonlinearity (e.g. Mal and PEP) observed with the ¹²C calibration lines, which may be caused by instrumental variations and/or ion suppression effects, are effectively corrected by the use of U-¹³C-labeled IS. This demonstrates the principle of isotope dilution, i.e. the ratio between the U-¹³C-labeled IS and the unlabeled analyte is much less sensitive

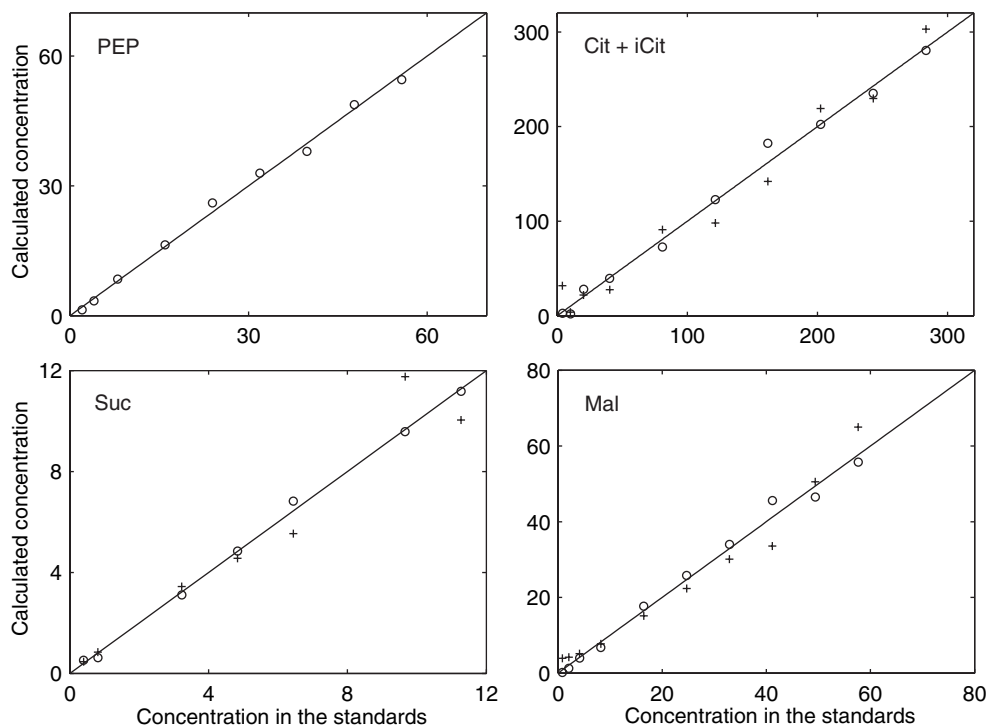


Figure 3.1 Parity plots for selected metabolites. Known concentrations in the standards (x-axis, in μM) are plotted against calculated concentrations (y-axis, in μM) according to the IS-based calibration line (o) and ^{12}C -based calibration line (+).

to adverse experimental and instrumental conditions than the absolute peak areas.

The slope of each IS-based calibration line equals the reciprocal of the corresponding $\text{U-}^{13}\text{C}$ -labeled metabolite concentration in the labeled cell extracts and should in principle be independent of possible instrumental variations and should therefore show no variation between different analysis runs. To confirm this, a single set of calibration standards, to which labeled extracts were added, was subjected to four independent calibration runs (see Table 3.3). In the last two calibration runs a decrease of both the ^{12}C and $\text{U-}^{13}\text{C}$ peak areas was observed for all metabolites, indicating either loss of sensitivity of the LC-ESI-MS/MS or degradation of metabolites. This leads to large changes in the slopes of ^{12}C -based calibration lines (for F6P e.g. the estimated slopes in the four runs, in the same sequence as Table 3.3, are 207.6, 120.6, 11.0 and 20.4) but hardly effects the IS-based calibration, demonstrating the robustness of the applied isotope dilution methodology.

Metabolite analysis in steady state and transient samples IDMS analysis was applied to the steady state samples from a chemostat culture of *S. cerevisiae* and transient samples when the chemostat was perturbed by a pulse of concentrated glucose. Absolute concentrations of intracellular metabolites were obtained from the peak area ratios of ^{12}C - and $\text{U-}^{13}\text{C}$ -labeled metabolites in the samples and the IS-

Table 3.3 Slopes and their 95% confidence intervals of selected IS-based calibration lines in four LC-ESI-MS/MS runs

| | G6P | F6P | F1,6bP | PEP | α KG | Fum |
|-------|-------------|-----------|-----------|-----------|-------------|-----------|
| Run 1 | 0.042±0.002 | 0.30±0.02 | 0.24±0.01 | 0.59±0.03 | 0.15±0.01 | 0.32±0.02 |
| Run 2 | 0.042±0.002 | 0.32±0.02 | 0.24±0.01 | 0.60±0.05 | 0.16±0.01 | 0.31±0.01 |
| Run 3 | 0.042±0.002 | 0.28±0.01 | 0.25±0.00 | 0.59±0.03 | 0.16±0.01 | 0.32±0.02 |
| Run 4 | 0.042±0.002 | 0.29±0.03 | 0.24±0.01 | 0.63±0.02 | 0.15±0.01 | 0.33±0.02 |

based calibration lines. This was possible because the labeled cell extract was added to the calibration standards and to the samples to the same final concentration (i.e. diluted 5 times). Due to the high physical-chemical resemblance of the U-¹³C-labeled IS and the unlabeled analyte, metabolite degradation during sample preparation, as well as ion suppression effects and variations in the sample analysis were corrected for. It should, however, be mentioned that addition of IS cannot correct for cell leakage during quenching and centrifugation as well as incomplete extraction of metabolites from the cells. For *S. cerevisiae*, it has been shown that the sample preparation procedure as applied in this study effectively preserves cell integrity by quenching with cold methanol and leads to complete release of the metabolites by extraction with boiling ethanol (Gonzalez et al., 1997).

A number of metabolites (e.g. G1P and T6P) are present in the labeled extracts in extremely low concentrations and therefore cannot be properly detected in the current LC-ESI-MS/MS setup. Therefore the concentration of these metabolites in samples cannot be determined by means of the isotope dilution method. This might, however, be circumvented by e.g. adding a larger amount of labeled extract to samples and standards, increasing sample injection volumes or applying different fed-batch cultivation strategies for the production of labeled extract, leading to different concentrations of intracellular metabolites.

The average steady state concentrations, determined from 16 steady state samples, are given in Table 3.4 for each metabolite. The absolute concentrations of glycolytic intermediates correspond very well with those reported in (Mashego et al., 2004), determined by conventional LC-MS/MS analysis. The temporal profiles of several representative intracellular metabolites during the transient after a glucose pulse are plotted in Fig. 3.2. The dynamics of the glycolytic intermediates is comparable to published data (Mashego et al., 2004), showing increase in the hexose phosphates concentrations and decreases in the concentration of 2PG, 3PG and PEP. The concentrations of TCA cycle metabolites are much less perturbed than the glycolytic intermediates, since major part of the increased carbon influx is redirected to ethanol instead of entering the TCA cycle.

Precision of isotope dilution LC-ESI-MS/MS To evaluate the precision of IDMS analysis, the relative standard errors (RSE) of each metabolite using IDMS and conventional ¹²C-based LC-ESI-MS/MS analysis were compared. For each metabolite the RSE of the two analytical methods was evaluated by calculating the RSE of each duplicate measurement of all 42 samples (steady state and transient) from the *S. cerevisiae* fermentation: for IDMS, 42 RSEs of peak area ratios were calculated;

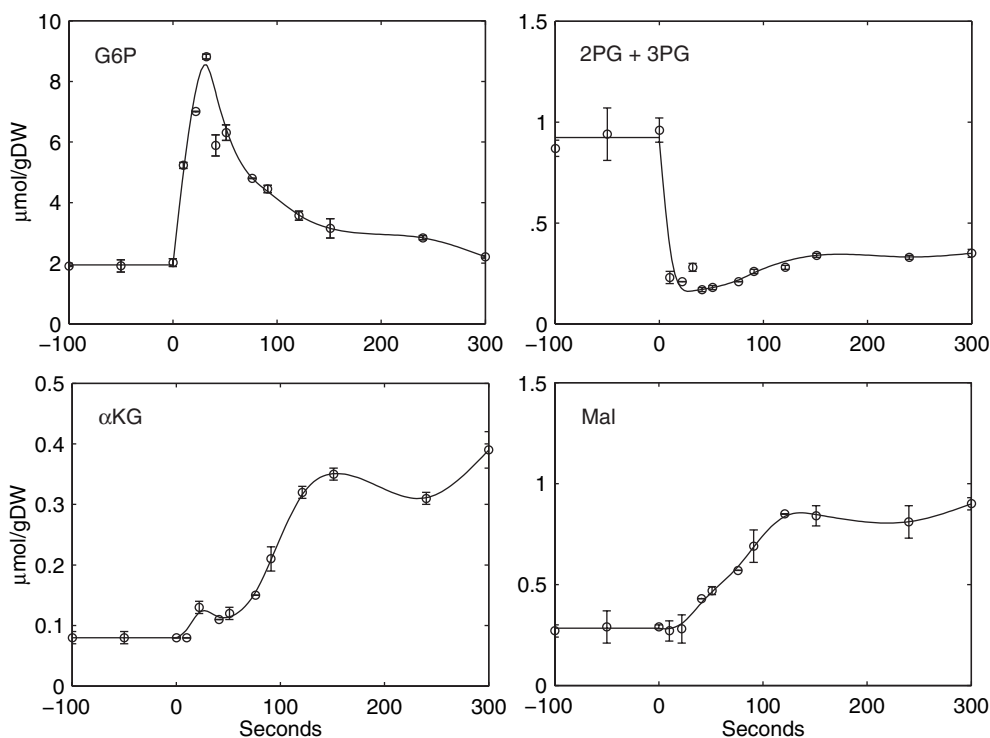


Figure 3.2 Intracellular metabolite concentrations ($\mu\text{mol/gDW}$) following a glucose pulse to a chemostat culture of *S. cerevisiae*.

Table 3.4 Steady state intracellular metabolite concentrations ($\mu\text{mol/gDW}$) in a chemostat culture of *S. cerevisiae*^a

| Metabolite | Concentration | Metabolite | Concentration |
|------------|-----------------|-------------------|-----------------|
| G6P | 2.11 ± 0.05 | Cit+iCit | 5.57 ± 0.11 |
| F6P | 0.38 ± 0.02 | αKG | 0.09 ± 0.00 |
| F1,6bP | 0.20 ± 0.03 | Suc | 0.05 ± 0.00 |
| 2PG+3PG | 1.33 ± 0.02 | Fum | 0.06 ± 0.00 |
| PEP | 1.05 ± 0.05 | Mal | 0.28 ± 0.02 |

^a Concentrations and 95% confidence intervals were determined from duplicate measurements of 16 steady state samples.

for the conventional ^{12}C -based analysis, 84 RSEs of peak areas of both the ^{12}C - and $\text{U-}^{13}\text{C}$ -labeled metabolites were obtained. By using duplicate measurements of the same sample, error sources other than the MS-analysis are excluded. Moreover, the fairly large number of the calculated RSEs ensures that the obtained distribution is representative for both analytical methods. The average RSEs for each metabolite obtained from the two analytical methods are given in Table 3.5, which demonstrate that isotope dilution leads to much smaller RSEs compared to the ^{12}C -based method and therefore provides metabolomic data with a higher precision.

Table 3.5 Precision of the IDMS analysis evaluated by relative standard errors (RSE)

| Metabolite | RSE | | |
|-------------|--|-------------------|--|
| | steady state samples by IDMS ^a | IDMS ^b | ¹² C-based LC-ESI-MS/MS ^c |
| G6P | 4.9 | 3.7 | 13.3 |
| F6P | 7.8 | 11.3 | 15.7 |
| F1,6bP | 10.4 | 3.7 | 9.6 |
| 2PG+3PG | 4.2 | 6.0 | 10.2 |
| PEP | 9.5 | 7.5 | 13.5 |
| Cit+iCit | 3.7 | 3.4 | 5.4 |
| α KG | 11.1 | 6.0 | 6.1 |
| Suc | 20.1 | 12.3 | 14.8 |
| Fum | 13.1 | 8.2 | 13.7 |
| Mal | 10.3 | 7.4 | 14.5 |

^a RSE of the intracellular concentration measured in 16 steady state samples.

^b Average of RSEs from 42 duplicate area ratio measurements of the same steady state samples and transient samples.

^c Average of RSEs from 84 duplicate ¹²C and U-¹³C area measurements of the same steady state samples and transient samples.

The decrease in the average RSE of duplicate measurements indicates that the influence of variations during LC-ESI-MS/MS analysis is effectively reduced by isotope dilution. A number of error sources during sample preparation (e.g. metabolite degradation) can be effectively eliminated by isotope dilution as well. However, inaccuracies in sample weight determination as well as errors in added volumes of labeled extracts could still influence the accuracy of the final result. To evaluate the possible contribution of these remaining error sources, we calculated the RSEs of 16 individually processed steady state samples of *S. cerevisiae* (i.e. containing all experimental error sources) and compared those in Table 3.5 with the average RSE of the duplicate measurements of single samples by IDMS. It should be mentioned that due to the limited population of steady state samples, the calculated RSEs of steady state concentrations might not be precise. Table 3.5 shows that more variations are present in steady state concentrations than could be accounted for by errors in IDMS analysis alone (except F6P and 2+3PG), indicating that additional errors might be involved in the sample preparation. It can however not be excluded that variation in steady state metabolite concentrations reflects partially the true fluctuations of intracellular metabolites in cells present in the fermentor occurring during the time period that the samples were taken.

In conclusion, fully labeled cell extracts, obtained by fed-batch cultivation of *S. cerevisiae* on fully U-¹³C-labeled substrates, were used as internal standards in the IDMS analysis of low molecular weight metabolites, present typically in μ M range in samples. We have demonstrated, with the analysis of glycolytic and TCA cycle intermediates in yeast, that the IDMS method presented here significantly improves the precision of metabolite concentration measurements (by about a factor 2, see

the last two columns in Table 3.5) and the linearity of calibration lines compared to conventional LC-MS/MS analysis. Method-development effort can be largely reduced as co-extraction of the labeled extract together with the samples eliminates the need for laborious recovery checks. Extension of the presented method to other mass spectrometric analysis can be readily made as any possible metabolite degradation (e.g. during derivatization steps needed for GC-MS) can be accounted for by labeled internal standards. IDMS with labeled extracts as internal standards thus enables a robust quantification of intracellular metabolite concentrations with high precision and holds great potential in future metabolomics research.

References

- Annesley TM. 2003. Ion suppression in mass spectrometry. *Clin Chem* 49: 1041–4.
- Baillie TA. 1981. The use of stable isotopes in pharmacological research. *Pharmacol Rev* 33: 81–132.
- Buchholz A, Hurlebaus J, Wandrey C, Takors R. 2002. Metabolomics: quantification of intracellular metabolite dynamics. *Biomol Eng* 19: 5–15.
- Buchholz A, Takors R, Wandrey C. 2001. Quantification of intracellular metabolites in *Escherichia coli* K12 using liquid chromatographic-electrospray ionization tandem mass spectrometric techniques. *Anal Biochem* 295: 129–37.
- de Koning W, van Dam K. 1992. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Anal Biochem* 204: 118–23.
- Delneri D, Brancia FL, Oliver SG. 2001. Towards a truly integrative biology through the functional genomics of yeast. *Curr Opin Biotechnol* 12: 87–91.
- Evans J, Wang TC, Heyes MP, Markey SP. 2002. LC/MS analysis of NAD biosynthesis using stable isotope pyridine precursors. *Anal Biochem* 306: 197–203.
- Fiehn O. 2002. Metabolomics - the link between genotypes and phenotypes. *Plant Mol Biol* 48: 155–71.
- Gonzalez B, Francois J, Renaud M. 1997. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast* 13: 1347–55.
- Grey L, Nguyen B, Yang P. 2002. Liquid chromatography-electrospray ionization isotope dilution mass spectrometry analysis of paraquat and diquat using conventional and multilayer solid-phase extraction cartridges. *J Chromatogr A* 958: 25–33.
- Hellerstein MK, Neese RA. 1999. Mass isotopomer distribution analysis at eight years: theoretical, analytic, and experimental considerations. *Am J Physiol* 276: E1146–70.
- Katz JJ, Crespi HL. 1966. Deuterated organisms: cultivation and uses. *Science* 151: 1187–94.
- Kollman VH, Hanners JL, Hutson JY, Whaley TW, Ott DG, Gregg CT. 1973. Large-scale photosynthetic production of carbon-13 labeled sugars: the tobacco leaf system. *Biochem Biophys Res Commun* 50: 826–31.
- Kuklenyik Z, Ashley DL, Calafat AM. 2002. Quantitative detection of trichloroacetic acid in human urine using isotope dilution high-performance liquid chromatography-electrospray ionization tandem mass spectrometry. *Anal Chem* 74: 2058–63.

- Lange HC, Eman M, van Zuijlen G, Visser D, van Dam JC, Frank J, de Mattos MJ, Heijnen JJ. 2001. Improved rapid sampling for *in vivo* kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 75: 406–15.
- Magni F, Pereira S, Leoni M, Grisenti G, Galli Kienle M. 2001. Quantitation of cyclosporin A in whole blood by liquid chromatography/stable isotope dilution electrospray ionization mass spectrometry. *J Mass Spectrom* 36: 670–6.
- Mashego MR, Wu L, Van Dam JC, Ras C, Vinke JL, Van Winden WA, Van Gulik WM, Heijnen JJ. 2004. MIRACLE: mass isotopomer ratio analysis of U-¹³C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. *Biotechnol Bioeng* 85: 620–8.
- Matwiyoff NA, Ott DG. 1973. Stable isotope tracers in the life sciences and medicine. *Science* 181: 1125–33.
- Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG. 2001. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* 19: 45–50.
- Rizzi M, Baltes M, Theobald U, Reuss M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol Bioeng* 55: 592–608.
- Sojo LE, Lum G, Chee P. 2003. Internal standard signal suppression by co-eluting analyte in isotope dilution LC-ESI-MS. *Analyst* 128: 51–4.
- Theobald U, Mailinger W, Baltes M, Reuss M, Rizzi M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng* 55: 305–16.
- Tweeddale H, Notley-McRobb L, Ferenci T. 1998. Effect of slow growth on metabolism of *Escherichia coli*, as revealed by global metabolite pool (“metabolome”) analysis. *J Bacteriol* 180: 5109–16.
- van Dam JC, Eman M, Frank J, Lange HC, van Dedem GW, Heijnen JJ. 2002. Analysis of glycolytic intermediates in *Saccharomyces cerevisiae* using anion exchange chromatography and electrospray ionization with tandem mass spectrometric detection. *Anal Chim Acta* 460: 209–18.
- Verduyn C, Postma E, Scheffers WA, Van Dijken JP. 1992. Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 8: 501–17.
- Visser D. 2002. Measuring and modeling *in vivo* kinetics of primary metabolism, PhD Thesis. Delft University of Technology, Delft.

Chapter 4

Lin-Log Kinetics for Control Parameter Estimation

Abstract The control properties of biochemical pathways can be described by control coefficients and elasticities, as defined in the framework of Metabolic Control Analysis (MCA). The determination of these parameters using the traditional MCA relationships is however limited by experimental difficulties (e.g. realizing and measuring small changes in biological systems) and lack of appropriate mathematical procedures (e.g. when the more practical large changes are made). In this chapter, the lin-log kinetics is proposed to avoid the above-mentioned problems and is applied to estimate control parameters from measurements obtained in steady state experiments. The lin-log approach employs approximative linear-logarithmic kinetics parameterized by elasticities and provides analytical solutions for fluxes and metabolite concentrations when large changes are made. Published flux and metabolite concentration data are used, obtained from a reconstructed section of glycolysis converting 3-phosphoglycerate to pyruvate [Giersch, C. (1995) *Eur. J. Biochem.* 227, 194-201]. With the lin-log approach, all data from different experiments can be combined to give realistic elasticity and flux control coefficient estimates by linear regression. Despite the large changes, a good agreement of fluxes and metabolite concentrations is obtained between the measured and calculated values according to the lin-log model. Furthermore, it is shown that the lin-log approach allows a rigorous statistical evaluation to identify the optimal reference state and the optimal model structure assumption. In conclusion, the lin-log approach addresses practical problems encountered in the traditional MCA-based methods by introducing suitable nonlinear kinetics, with elasticities as kinetic parameters, thus providing a novel framework with improved procedures for the estimation of elasticities and control parameters from large perturbation experiments.

Published as: Wu L, Wang WM, van Gulik WM, Heijnen JJ. 2004. A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. *Eur J Biochem* 271: 3348-59.

4.1 Introduction

Cellular metabolism is complex and interconnected by nature, and the quantitative understanding of it requires mathematical modeling tools (Wiechert, 2002). Metabolic control analysis (MCA) (Heinrich and Rapoport, 1974; Kacser and Burns, 1973) provides such a mathematical framework that quantifies the changes in metabolic variables (such as fluxes and concentrations) as a result of changes in system parameters (such as enzyme levels and metabolite concentrations) through sensitivity analysis. In the context of a rational redesign of metabolic networks (Bailey, 1991; Nielsen, 1998), the changes of fluxes by changing the levels of pathway enzymes is of particular interest. In MCA this is quantified by the flux control coefficient (FCC) C^J , defined as the scaled infinitesimal change of flux J_i in response to an infinitesimal change in enzyme level e_j around a reference (quasi-) steady-state (denoted by 0) (Fell, 1997):

$$C_{ij}^{J0} = \frac{e_j^0}{J_i^0} \left(\frac{dJ_i}{de_j} \right)^0 \quad (4.1)$$

Control coefficient represents a global property of the metabolic system, which is in turn determined by the kinetic property of individual enzymes, parameterized by their elasticities ε towards a dependent metabolite concentration x_k (which cannot be freely manipulated) or an independent metabolite concentration c_l (which can be freely manipulated):

$$\varepsilon_{ik}^{x0} = \frac{x_k^0}{v_i^0} \left(\frac{\partial v_i}{\partial x_k} \right)^0, \quad \varepsilon_{il}^{c0} = \frac{c_l^0}{v_i^0} \left(\frac{\partial v_i}{\partial c_l} \right)^0 \quad (4.2)$$

Since the introduction of MCA in the 1970's, extensive research has been devoted to the experimental determination of FCCs (Fell, 1997). These can be calculated directly according to their definitions, by introducing small changes in enzyme activities close to a reference state and measuring the resulting small changes in fluxes (i.e. the 'direct' method). Alternatively, FCCs can be calculated from elasticities (the 'indirect' method): small changes are introduced in either enzyme levels or independent metabolite concentrations around a reference state and the resulting small changes in fluxes and metabolite concentrations measured; these measurements are subsequently used in Eq. 4.2 to obtain the elasticities. According to the definition of MCA parameters, small perturbations of, e.g. the enzyme activity, must be realized accurately in a biological system and the resulting response of fluxes and metabolite concentrations, usually minor, must be recorded with high precision. This is experimentally demanding and the attainable precision often fails to yield reliable results (Giersch, 1995; Pettersson, 1996).

In practice, a series of larger perturbations has to be made to produce measurable effects. However, the MCA framework cannot directly deal with the resulting nonlinear responses, as it is only valid in the vicinity of the reference state. The infinitesimal perturbation required by MCA has to be approximated numerically by first fitting the data with a nonlinear (e.g. polynomial) function (Small, 1993), followed by determination of the slope at the reference state. A large number of (quasi) steady-state measurements is required to ensure a good fit and a reliable slope at the chosen reference state. The fitting is inevitably subject to underlying assumptions of

the nonlinear function used and accompanied by the loss of information contained in individual data points. Moreover, this also gives rise to nontrivial statistical treatments, as attempted previously (Ainscow and Brand, 1998; Ehlde and Zacchi, 1996). Altogether it leads to a poor utilization of the available information obtained via such extensive experimental efforts.

Extensions of the MCA framework have been made to accommodate larger changes in e.g. fluxes, enzyme levels or metabolite concentrations (Heinrich and Schuster, 1996; Hofer and Heinrich, 1993; Small and Kacser, 1993a,b) and applied to simple metabolic systems (Simpson et al., 1998). The applicability of these extensions are rather limited, either due to difficulties in the determination of additional parameters (e.g. second-order elasticities) (Hofer and Heinrich, 1993), or to the mathematical complexity which arises when realistic metabolic networks are considered. Estimation of MCA parameters in practical conditions of larger changes is yet insufficiently addressed.

In the ideal situation of the current MCA approach, where small perturbations can be realized, the results obtained by different experimental methods (i.e. the ‘direct’ or ‘indirect’ method) or at different reference states have to be treated as separate datasets and provide elasticities and FCCs at the specific reference state only. At this moment it is not possible to combine different types of perturbations performed at different reference states within the same experimental setup to give one set of MCA parameter estimates.

In this chapter, an improvement on the mathematical approach to MCA parameter estimation is proposed and elucidated by an example. The new approach is based on a linear-logarithmic approximation of enzyme kinetics (lin-log kinetics) (Visser and Heijnen, 2003). In the lin-log kinetics the reaction rate is a nonlinear function of metabolite concentrations and proportional to enzyme levels. This allows a satisfactory approximation of large perturbations in both metabolite and enzyme levels, as often encountered in biological experiments. In addition, the lin-log approach provides general steady state solutions for flux and dependent metabolite concentrations, as a function of changes in enzyme levels and independent metabolite concentrations. The lin-log approach also allows to combine easily all types of experimental information (e.g. measurement of fluxes, enzyme levels and metabolite concentrations in different reference states) into one consistent set of MCA parameter estimates. Finally, in the lin-log kinetics the elasticities are present in a linear form and therefore can be estimated using simple linear regression. In this article, the lin-log approach is tested on data from a set of experiments published by Giersch (1995), in which different perturbation experiments were performed on a linear reconstituted (i.e. *in vitro*) pathway. A comparison between the lin-log approach and the multiple modulation method used by Giersch is made, based on the MCA parameters estimated from both methods.

4.2 Theory

Relations from the metabolic control analysis Consider a metabolic network consisting of m dependent metabolites with concentrations x_k ($k = 1, 2, \dots, m$), r independent metabolites with concentration c_l ($l = 1, 2, \dots, r$) and n enzymes

Table 4.1 Definition of MCA parameters

| | |
|--|--|
| Elasticity with respect to x_k | Flux control coefficient |
| $\varepsilon_{ik}^{x0} = \frac{x_k^0}{J_i^0} \left(\frac{\partial v_i}{\partial x_k} \right)^0$ | $C_{ij}^{J0} = \frac{e_j^0}{J_i^0} \left(\frac{dJ_i}{de_j} \right)^0$ |
| Elasticity with respect to e_j | Metabolite control coefficient |
| $\varepsilon_{ij}^{e0} = \frac{e_j^0}{J_i^0} \left(\frac{\partial v_i}{\partial e_j} \right)^0$ | $C_{kj}^{x0} = \frac{e_j^0}{x_k^0} \left(\frac{dx_k}{de_j} \right)^0$ |
| Elasticity with respect to c_l | Flux response coefficient |
| $\varepsilon_{il}^{c0} = \frac{c_l^0}{J_i^0} \left(\frac{\partial v_i}{\partial c_l} \right)^0$ | $R_{il}^{J0} = \frac{c_l^0}{J_i^0} \left(\frac{dJ_i}{dc_l} \right)^0$ |
| | Metabolite response coefficient |
| | $R_{kl}^{x0} = \frac{c_l^0}{x_k^0} \left(\frac{dx_k}{dc_l} \right)^0$ |

Table 4.2 MCA relations for a linear pathway

| | |
|--|--|
| $\mathbf{C}^{x0} = -(\mathbf{S}[\mathbf{J}^0] \mathbf{E}^{x0})^{-1} \cdot \mathbf{S} \cdot [\mathbf{J}^0]$ | $\mathbf{C}^{J0} = \mathbf{E}^{x0} \mathbf{C}^{x0} + \mathbf{i}$ |
| $\mathbf{R}^{x0} = \mathbf{C}^{x0} \mathbf{E}^{c0}$ | $\mathbf{R}^{J0} = \mathbf{C}^{J0} \mathbf{E}^{c0}$ |

with enzyme level e_i ($i = 1, 2, \dots, n$), each catalyzing a reaction with reaction rate v_i . Table 4.1 summarizes the definition of MCA parameters that follows from the metabolic network under consideration (Visser and Heijnen, 2002).

A set of relations between the global system properties (such as the control and response coefficients) and the elasticities can be derived following various mathematical procedures (Visser and Heijnen, 2002). The MCA relations for a linear pathway without conserved moieties, which will be dealt with further, are summarized in Table 4.2.

Lin-log kinetic approach The general dependency of a reaction rate v_i on enzyme levels, dependent and independent metabolite concentrations can be expressed, with a nonlinear function f_i , as:

$$v_i = f_i(\mathbf{e}, \mathbf{x}, \mathbf{c}) \quad (4.3)$$

In the lin-log kinetic formulation, the nonlinear function f_i in Eq. 4.3 is parameterized as

$$v_i = e_i(a_i + p_{i,1} \ln x_1 + p_{i,2} \ln x_2 + \dots + p_{i,m} \ln x_m + q_{i,1} \ln c_1 + q_{i,2} \ln c_2 + \dots + q_{i,r} \ln c_r) \quad (4.4)$$

The reaction rate v_i in Eq. 4.3 is assumed to be proportional to the enzyme level e_i and a linear combination of logarithms of dependent and independent metabolite concentrations.

Approximative kinetics involving logarithmic concentration terms has its origin in the concept that the rate of a chemical reaction is proportional to the reaction affinity A when the rate is close to equilibrium (Onsager, 1931):

$$v = LA$$

For e.g. a reaction $x_1 \rightleftharpoons x_2$ the affinity is given by:

$$A = RT \ln \left(K_{eq} \frac{x_1}{x_2} \right)$$

Rottenberg (1973) and van der Meer et al. (1980) have shown that for an enzyme-catalyzed reaction far from equilibrium, the reaction rate can be linear (but not proportional) with respect to reaction affinity in a wide range of reaction affinities, i.e.

$$v = L^\#(A - A^\#) = (L^\# RT \ln K_{eq} - L^\# A^\#) + L^\# RT \ln x_1 - L^\# RT \ln x_2 \quad (4.5)$$

where $L^\#$ and $A^\#$ are kinetic coefficients solely determined by the enzyme kinetic parameters. An extensive description of this concept is given in Westerhoff and van Dam (1987).

Two points deserve attention regarding Eq. 4.5. First, the rate of an enzyme-catalyzed reaction is not only influenced by the reactants involved (i.e. x_1 and x_2), but in general also by allosteric effectors. These effectors increase or decrease the reaction rate without contributing to the reaction affinity. To take such kinetic effects into account, Westerhoff and van Dam (1987) had $L^\#$ also depend on the concentration of these effectors. Indeed, this is the very basis of their Mosaic Non Equilibrium Thermodynamic approach which, in contrast to classical non equilibrium thermodynamics, can be used to address mechanisms (see e.g. Hellingwerf et al. 1979; Westerhoff et al. 1981, 1982). Later, Nielsen (1997) proposed to extend Eq. 4.5 by including logarithmic concentration terms of allosteric effectors using empirical kinetic coefficients. A second point is that the kinetic coefficients for the logarithmic concentrations in Eq. 4.5 (i.e. $L^\# RT$ for $\ln x_1$ and $-L^\# RT$ for $\ln x_2$) are related to each other, having the same absolute value but opposite signs. The dependency between the kinetic coefficients has its origin in using the reaction affinity A as the sole variable in the reaction kinetics. This seems unnecessarily restrictive. Westerhoff and Van Dam introduced the so-called ‘asymmetry coefficients’ or ‘force-stoichiometries’ (γ), which allow the reaction rate to be differentially sensitive to the logarithms of the concentrations of the various reactants (van Dam et al., 1981; Westerhoff and van Dam, 1987). Likewise, we adopt independent, empirical coefficients in Eq. 4.4, not only for each allosteric effector, but also for individual reactants.

A reference state can be introduced in Eq. 4.4, with enzyme level e_i^0 , dependent metabolite concentration x_k^0 , independent metabolite concentration c_l^0 and steady state flux J_i^0 . It follows from Eq. 4.4 that:

$$J_i^0 = e_i^0 (a_i + p_{i,1} \ln x_1^0 + p_{i,2} \ln x_2^0 + \dots + p_{i,m} \ln x_m^0 + q_{i,1} \ln c_1^0 + q_{i,2} \ln c_2^0 + \dots + q_{i,r} \ln c_r^0) \quad (4.6)$$

Using the definition of elasticity in Eq. 4.2, one obtains for reaction i the elasticities in the reference state for the k^{th} dependent metabolite ($\varepsilon_{ik}^{x_0}$) and the l^{th} independent

metabolite (ε_{il}^{c0}):

$$\varepsilon_{ik}^{x0} = \frac{p_{i,k} e_i^0}{J_i^0}, \quad \varepsilon_{il}^{c0} = \frac{q_{i,l} e_i^0}{J_i^0} \quad (4.7)$$

Combining Eq. 4.4, 4.6 and 4.7, the kinetic parameters a_i , $p_{i,k}$ and $q_{i,l}$ can be eliminated to give the reference based lin-log kinetic equation:

$$\begin{aligned} \frac{v_i}{J_i^0} = \frac{e_i}{e_i^0} & \left(1 + \varepsilon_{i,1}^{x0} \ln \frac{x_1}{x_1^0} + \varepsilon_{i,2}^{x0} \ln \frac{x_2}{x_2^0} + \cdots + \varepsilon_{i,m}^{x0} \ln \frac{x_m}{x_m^0} \right. \\ & \left. + \varepsilon_{i,1}^{c0} \ln \frac{c_1}{c_1^0} + \varepsilon_{i,2}^{c0} \ln \frac{c_2}{c_2^0} + \cdots + \varepsilon_{i,r}^{c0} \ln \frac{c_r}{c_r^0} \right) \end{aligned} \quad (4.8)$$

For the whole metabolic network considered, Eq. 4.8 can be written in a matrix form:

$$\frac{\mathbf{v}}{\mathbf{J}^0} = \begin{bmatrix} \mathbf{e} \\ \mathbf{e}^0 \end{bmatrix} \left(\mathbf{i} + \mathbf{E}^{x0} \ln \frac{\mathbf{x}}{\mathbf{x}^0} + \mathbf{E}^{c0} \ln \frac{\mathbf{c}}{\mathbf{c}^0} \right) \quad (4.9)$$

with \mathbf{i} being the unit vector and \mathbf{E}^{x0} and \mathbf{E}^{c0} the elasticity matrix for dependent metabolites (\mathbf{x}) and independent metabolites (\mathbf{c}) respectively. The logarithm-based nonlinear dependencies in Eq. 4.9 allow the approximation of reaction rates over a broad range of enzyme levels and metabolite concentrations, which is illustrated for Michaelis-Menten type kinetics in Fig. 4.1 (the conversion of Michaelis-Menten kinetics to the corresponding lin-log kinetics and the linear kinetics is given in the Appendix A). The approximative performance of lin-log kinetics has been further demonstrated successfully for an *in silico* pathway involving highly nonlinear kinetics, conserved moieties and large changes in fluxes and metabolite concentrations (Visser and Heijnen, 2003). It is therefore possible to apply lin-log kinetic formulation to measurements obtained from large perturbation experiments (i.e. \mathbf{v}/\mathbf{J}^0 , \mathbf{e}/\mathbf{e}^0 , \mathbf{x}/\mathbf{x}^0 and \mathbf{c}/\mathbf{c}^0 , with which \mathbf{E}^{x0} and \mathbf{E}^{c0} can be estimated from Eq. 4.9 by linear regression.

Having obtained the elasticity parameters for Eq. 4.9, a full dynamic model of the whole metabolic network can be set up. Such a model allows in principle the calculation of control coefficients also under dynamic conditions. Moreover, the lin-log formulation enables the analytical solution of the steady state mass balances over the metabolic network (Visser and Heijnen, 2003), which gives rise to the following relationships:

$$\mathbf{C}^{\mathbf{J}^0} \left(\frac{\mathbf{v}}{\mathbf{J}^0} \frac{\mathbf{e}^0}{\mathbf{e}} \right) = \mathbf{i} + \mathbf{R}^{\mathbf{J}^0} \ln \frac{\mathbf{c}}{\mathbf{c}^0} \quad (4.10)$$

$$\ln \frac{\mathbf{x}}{\mathbf{x}^0} = -\mathbf{C}^{x0} \left(\frac{\mathbf{v}}{\mathbf{J}^0} \frac{\mathbf{e}^0}{\mathbf{e}} \right) + \mathbf{C}^{x0} \ln \frac{\mathbf{c}}{\mathbf{c}^0} \quad (4.11)$$

When the reference state elasticities are known, the control and response coefficients in Eq. 4.10 and 4.11 can be calculated according to MCA relationships given in Table 4.2. The above two equations therefore provide the steady state fluxes (\mathbf{J}^0) and dependent metabolite concentrations (\mathbf{x}), as a function of the (reference) enzyme levels (\mathbf{e} , \mathbf{e}^0) and the (reference) independent metabolite concentrations (\mathbf{c} , \mathbf{c}^0).

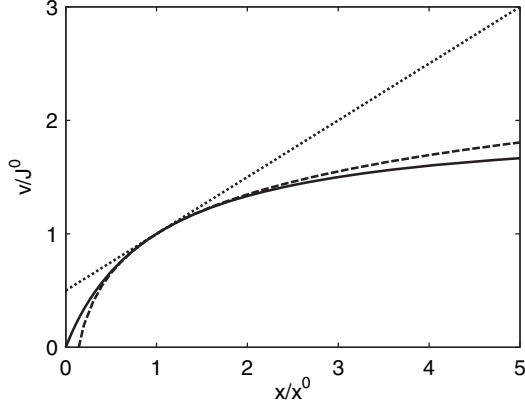


Figure 4.1 Approximation of the Michaelis-Menten kinetics (—) by linearization (\cdots) and lin-log kinetics (- - -), taking $e/e^0 = 1$, $\varepsilon_x^0 = 0.5$.

In perturbation experiments where the enzyme levels are altered and the independent metabolite concentrations remain constant, Eq. 4.10 reduces to

$$\mathbf{C}^{\mathbf{J}^0} \left(\frac{\mathbf{v}}{\mathbf{J}^0} \frac{\mathbf{e}^0}{\mathbf{e}} \right) = \mathbf{i} \quad (4.12)$$

With Eq. 4.12 FCCs in the reference state can be directly estimated, using measured large perturbations in \mathbf{v}/\mathbf{J}^0 and \mathbf{e}/\mathbf{e}^0 .

It has been long realized that elasticities are related to the kinetics of individual enzymes; in fact, when the enzyme kinetics are known, the elasticity at the reference state can be calculated straight-forward from the assumed enzyme kinetics. It is also known that the elasticities depend on the chosen reference state. However, the relationship between elasticities at different reference states cannot be calculated within the conventional MCA framework. In the lin-log formulation, elasticities serve directly as kinetic parameters in the nonlinear kinetic approximation (Eq. 4.9, see also Fig. 4.1). Analogous to the nonlinear hyperbolic types of kinetics from which elasticities can be calculated, it is possible to calculate elasticities from lin-log kinetics at arbitrary reference states ($\varepsilon_{ik}^{x'}$), based on elasticity at a chosen reference state ($\varepsilon_{ik}^{x_0}$). From Eq. 4.7 it follows that

$$\varepsilon_{ik}^{x'} = \frac{p_{i,k} e'_i}{J'_i} \quad (4.13)$$

Combining Eq. 4.7 and 4.13 by eliminating $p_{i,k}$, one obtains the relationship between the elasticity $\varepsilon_{ik}^{x_0}$ and $\varepsilon_{ik}^{x'}$:

$$\varepsilon_{ik}^{x'} = \varepsilon_{ik}^{x_0} \left[\frac{J_i^0 e'_i}{e_i^0 J'_i} \right] \quad (4.14)$$

Similar relationships hold for the elasticities of independent metabolites.

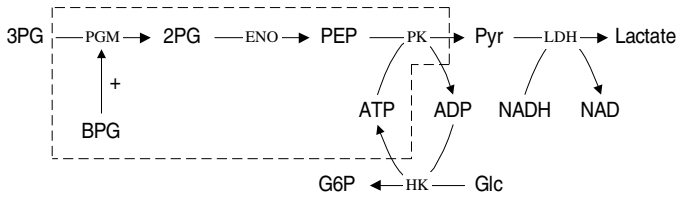


Figure 4.2 The reconstituted pathway (Giersch, 1995). The assumed stimulatory effects of BPG on PGM and ADP on PK is indicated by ‘+’ signs. The system boundary is indicated by the dashed lines. LDH: lactate dehydrogenase.

4.3 Methods

Experimental approach of Giersch The *in vitro* reconstituted pathway used by Giersch (1995) is illustrated in Fig. 4.2. In brief, the pathway consists of three glycolytic pathway enzymes: PGM, ENO and PK, converting 3PG to pyruvate. 2PG and PEP are considered as dependent metabolites; 2,3-bisphosphoglycerate (BPG) and ADP are independent metabolites, which are assumed to modulate specifically the activity of PGM and PK, respectively. A quasi steady state was established using a sufficiently high concentration of 3PG, the continuous regeneration of ADP (by adding HK and glucose) and the continuous removal of pyruvate (by irreversibly converting pyruvate to lactate with lactate dehydrogenase and NADH). The linear decrease of NADH fluorescence and the minimal changes in metabolite concentrations after approximately 1 hour confirmed the establishment of a quasi steady state and the choice of the system boundary, as indicated by the dashed lines in Fig. 4.2.

Two sets of quasi steady-state experiments were performed and the results, plotted in Fig. 3 and Fig. 4 in (Giersch, 1995), are summarized in Table 4.3. In the first set of experiments (Table 4.3, nr. 2 to 10), the concentrations of BPG and ADP were varied separately while the enzyme levels were constant; in the second set of experiments (Table 4.3, nr 11 to 19), the three enzyme levels were altered one by one and the concentrations of ADP and BPG were constant. Flux and metabolite concentrations were measured in all experiments. It was attempted to establish the same reference state within either sets of experiments. However, the same initial concentrations of reagents didn’t always result in the same quasi steady state, as can be seen in Table 4.3. In addition, Giersch used experiment 1 to demonstrate the establishment of a quasi steady state.

Elasticity estimation using the lin-log approach Prior to the elasticity estimation, an assumption about the pathway connectivity has to be made. Connectivity denotes which metabolites affect which reaction rates, hence, it determines how many elasticity parameters have to be estimated. Two types of connectivities were recognized by Giersch: for the mass-action type of connectivity, it is assumed that each reaction rate is affected only by its direct substrate(s) and product(s) and that the independent metabolites are specific, i.e. BPG affects only PGM and ADP affects only PK; for the maximum connectivity, every metabolite may affect every reaction rate. Since the reconstituted pathway contains 3 enzymes, 2 dependent metabolites and 2 independent metabolites, there are 12 unknown elasticities when maximum connectivity is assumed. In the case of mass-action connectivity, the

Table 4.3 Experimental data obtained by perturbation experiments (Giersch, 1995)

| Exp. nr. | Parameter changed ^a | PGM U/l | ENO U/l | PK U/l | 2PG μM | PEP μM | BPG μM | ADP μM | Flux $\mu\text{M/h}$ |
|-----------------|--------------------------------|---------|---------|--------|-------------------|-------------------|-------------------|-------------------|----------------------|
| 1 | | 31 | 5.6 | 46 | 170 | 50 | 13.2 | 107 | 157 |
| 2 | ADP | 25 | 12 | 30 | 60 | 103 | 13.2 | 61 | 87 |
| 3 | (61 – 130 μM) | 25 | 12 | 30 | 57 | 91 | 13.2 | 83 | 114 |
| 4 ^b | | 25 | 12 | 30 | 55 | 83 | 13.2 | 107 | 128 |
| 5 | | 25 | 12 | 30 | 54 | 76 | 13.2 | 130 | 136 |
| 6 | | 25 | 12 | 30 | 40 | 60 | 8.5 | 107 | 111 |
| 7 | BPG | 25 | 12 | 30 | 46 | 70 | 10.8 | 107 | 118 |
| 8 ^b | (8.5 – 18 μM) | 25 | 12 | 30 | 55 | 81 | 13.2 | 107 | 122 |
| 9 | | 25 | 12 | 30 | 59 | 89 | 15.6 | 107 | 130 |
| 10 | | 25 | 12 | 30 | 63 | 95 | 18 | 107 | 133 |
| 11 | PGM | 36.8 | 6.5 | 34 | 161 | 81 | 13.2 | 107 | 108 |
| 12 ^c | (36.8 – 53.2 U/l) | 45 | 6.5 | 34 | 178 | 82 | 13.2 | 107 | 108 |
| 13 | | 53.2 | 6.5 | 34 | 204 | 81 | 13.2 | 107 | 108 |
| 14 | ENO | 45 | 5.3 | 34 | 206 | 75 | 13.2 | 107 | 106 |
| 15 ^c | (5.3 – 7.7 U/l) | 45 | 6.5 | 34 | 195 | 87 | 13.2 | 107 | 108 |
| 16 | | 45 | 7.7 | 34 | 180 | 92 | 13.2 | 107 | 112 |
| 17 | PK | 45 | 6.5 | 30.9 | 192 | 91 | 13.2 | 107 | 96 |
| 18 ^c | (30.9 – 37.1 U/l) | 45 | 6.5 | 34 | 192 | 84 | 13.2 | 107 | 106 |
| 19 | | 45 | 6.5 | 37.1 | 196 | 79 | 13.2 | 107 | 116 |

^a The extent of changes is indicated between the brackets.

^b Reference state chosen by Giersch for the first set of experiments

^c Reference state chosen by Giersch for the second set of experiments

number of unknown elasticities reduces to 6.

Giersch compared both connectivity assumptions for the second set of experiments but used only the mass-action connectivity for the first set of experiments (maximum connectivity cannot be used as this would result in an under-determined system of equations).

In this paper, we compared both connectivity assumptions using data from all 19 experiments. From 19 experiments one can be chosen as the reference state. For the remaining 18 experiments, Eq. 4.8 can be rearranged in the following form for the reaction catalyzed by the i^{th} enzyme

$$\left(\frac{v_i}{J_i^0} \frac{e_i^0}{e_i} \right) - 1 = \varepsilon_{i,1}^{x_0} \ln \left(\frac{x_1}{x_1^0} \right) + \varepsilon_{i,2}^{x_0} \ln \left(\frac{x_2}{x_2^0} \right) + \varepsilon_{i,1}^{c_0} \ln \left(\frac{c_1}{c_1^0} \right) + \varepsilon_{i,2}^{c_0} \ln \left(\frac{c_2}{c_2^0} \right) \quad (4.15)$$

where x_1 , x_2 , c_1 and c_2 stands for the concentration of 2PG, PEP, BPG and ADP, respectively. The three reactions thus lead to a total of 54 linear equations in the form of Eq. 4.15 for each chosen reference state. Entering the measured values of enzyme levels, fluxes and metabolite concentrations, the elasticities belonging to the chosen reference state were estimated from Eq. 4.15 by multiple linear regression (Matlab R13, The Mathworks Inc., Natick, MA, USA).

Since there are 19 experiments, 19 different reference states can be chosen, leading to 19 sets of elasticities. To find an optimal reference state for the whole dataset, the convenient and widely applied R^2 -statistics (also called the coefficient of determination) was used to evaluate the goodness of fit of the 19 linear regressions. Since two connectivity assumptions are involved with different numbers of elasticity parameters, R^2 is augmented to accommodate the resulting difference in the degrees of freedom according to

$$R_a^2 = 1 - (1 - R^2) \frac{N}{N - P} \quad (4.16)$$

where R_a^2 is the adjusted coefficient of determination, N is the number of observations and P is the number of elasticities parameters (Draper and Smith, 1998). The reference state with the highest R_a^2 -value is regarded as the optimal reference state. Details of the multiple linear regression and adjusted R^2 -statistics can be found in Appendix B.

FCC estimation using the lin-log approach In experiment 11 to 19, the enzyme levels were changed while the concentration of BPG and ADP remained constant (see Table 4.3). According to Eq. 4.12, FCC can be directly estimated from these experiments. For the pathway considered it holds

$$\frac{J^0}{J} = C_1^{J^0} \begin{pmatrix} e_1^0 \\ e_1 \end{pmatrix} + C_2^{J^0} \begin{pmatrix} e_2^0 \\ e_2 \end{pmatrix} + C_3^{J^0} \begin{pmatrix} e_3^0 \\ e_3 \end{pmatrix} \quad (4.17)$$

where e_1 , e_2 and e_3 represent the enzyme levels of PGM, ENO and PK, respectively. At a chosen reference state out of experiment 11-19, the measured enzyme levels and fluxes were entered into Eq. 4.17, leading to 8 equations linear in the FCCs. These were then estimated by multiple linear regression (Matlab r13).

Alternatively, eliminating $C_1^{J^0}$ from Eq. 4.17 by the summation theorem reduces one degree of freedom, which gives:

$$\frac{J^0}{J} - \frac{e_1^0}{e_1} = C_2^{J^0} \begin{pmatrix} e_2^0 \\ e_2 \end{pmatrix} - \frac{e_1^0}{e_1} + C_3^{J^0} \begin{pmatrix} e_3^0 \\ e_3 \end{pmatrix} - \frac{e_1^0}{e_1} \quad (4.18)$$

The two remaining FCCs were estimated as described above. $C_1^{J^0}$ was subsequently calculated from the summation theorem.

4.4 Results and discussion

Elasticity estimation using the lin-log approach For each connectivity assumption, 19 sets of elasticities were obtained from multiple linear regression, each has one out of the 19 experiments as a reference state. To find an optimal reference state for the whole dataset, we used the adjusted coefficient of determination (R_a^2) as a criterion. Fig. 4.3 shows that experiment 2 provides the highest R_a^2 value out of 19 regressions for both connectivity types and is hence the best reference state for the whole dataset. The corresponding set of elasticities at reference experiment 2 and their 95% confidence intervals are given in Table 4.4 under ‘mass-action connectivity’ and ‘maximum connectivity’.

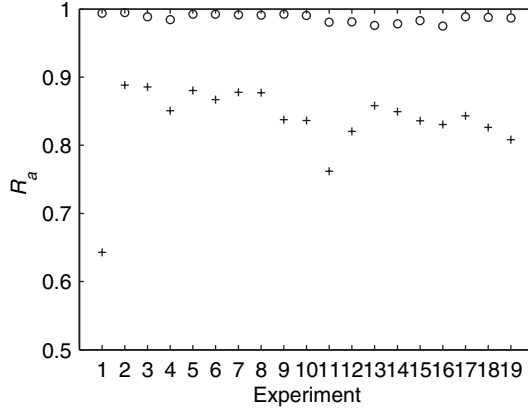


Figure 4.3 Adjusted coefficient of determination for 19 different reference states. +, Mass-action connectivity assumption; o, maximum connectivity assumption.

Table 4.4 Estimated elasticities at the optimal reference states^a

| Enzyme | Elasticity | Mass-action connectivity | | Maximum connectivity | | Optimized connectivity | |
|--------|---------------------------------------|--------------------------|----------------|----------------------|----------------|------------------------|----------------|
| PGM | $\varepsilon_{\text{PGM},2\text{PG}}$ | -0.24 | (± 0.14) | 0.56 | (± 0.05) | 0.38 | (± 0.03) |
| | $\varepsilon_{\text{PGM},\text{PEP}}$ | | | 1.16 | (± 0.23) | 0.89 | (± 0.04) |
| | $\varepsilon_{\text{PGM},\text{BPG}}$ | 0.37 | (± 0.86) | 1.45 | (± 0.25) | 1.05 | (± 0.13) |
| | $\varepsilon_{\text{PGM},\text{ADP}}$ | | | 0.17 | (± 0.12) | | |
| ENO | $\varepsilon_{\text{ENO},2\text{PG}}$ | 0.82 | (± 0.17) | 0.77 | (± 0.05) | 0.20 | (± 0.03) |
| | $\varepsilon_{\text{ENO},\text{PEP}}$ | -2.06 | (± 0.48) | -3.09 | (± 0.23) | -0.80 | (± 0.04) |
| | $\varepsilon_{\text{ENO},\text{BPG}}$ | - | | 1.83 | (± 0.25) | 0.48 | (± 0.13) |
| | $\varepsilon_{\text{ENO},\text{ADP}}$ | - | | -0.42 | (± 0.12) | - | |
| PK | $\varepsilon_{\text{PK},2\text{PG}}$ | - | | -0.27 | (± 0.05) | -0.21 | (± 0.02) |
| | $\varepsilon_{\text{PK},\text{PEP}}$ | 0.14 | (± 0.80) | -0.05 | (± 0.23) | - | |
| | $\varepsilon_{\text{PK},\text{BPG}}$ | - | | 0.56 | (± 0.25) | 0.46 | (± 0.12) |
| | $\varepsilon_{\text{PK},\text{ADP}}$ | 0.50 | (± 0.43) | 0.70 | (± 0.12) | 0.59 | (± 0.12) |

^a Elasticities which are set to zero, due to the connectivity assumption made, are indicated by '-'.
'-'

The goodness of fit reflected in the R_a^2 values can be further visualized in connection with the data in Table 4.3. Using the elasticities at reference experiment 2 with the measured enzyme levels and independent metabolite concentrations, we calculated the fluxes and dependent metabolite concentrations for each experiment with Eq. 4.10 and 4.11. In Fig. 4.4, the calculated values are set out against the measured ones in parity plots, which show that under mass-action connectivity assumption, all calculated fluxes and dependent metabolite concentrations deviate from the measured ones. Also, a rather low R_a^2 of 0.88 was found from the multiple linear regression with large confidence intervals of estimated elasticities (Table 4.4,

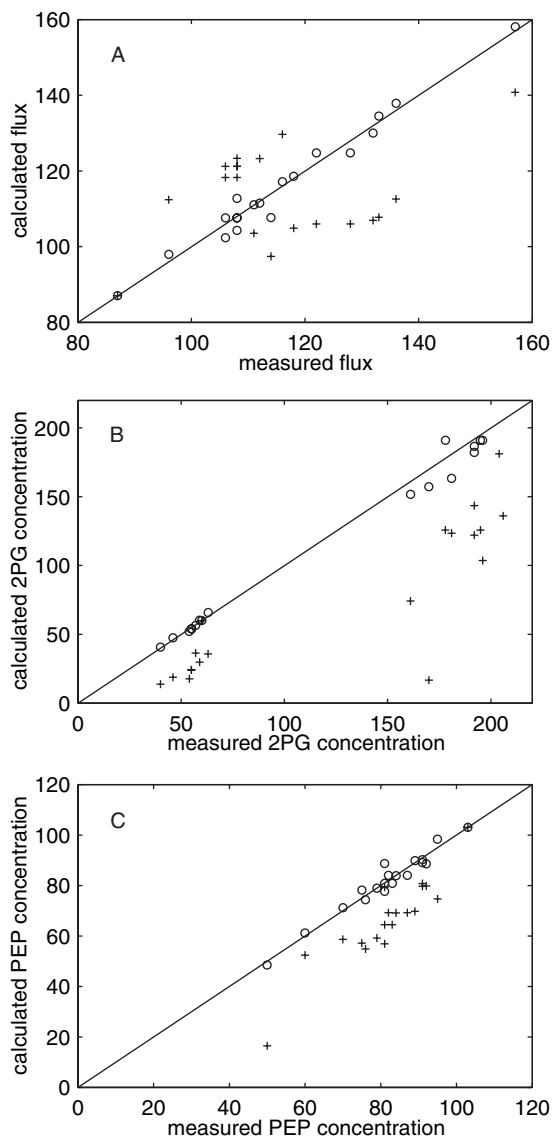


Figure 4.4 Parity plot of (A) measured and calculated fluxes (in $\mu\text{mol}/\text{l}/\text{h}$); (B) measured and calculated 2PG concentration (in μM); (C) measured and calculated PEP concentration (in μM). +, Mass-action connectivity assumption; o, maximum connectivity assumption.

mass-action connectivity).

The maximum connectivity assumption has an R_a^2 -value of 0.99 from the multiple linear regression. The calculated fluxes and dependent metabolite concentrations are all close to the measured values (Fig. 4.4). Since the degree of freedom has been taken into account in the evaluation of R_a^2 , the significantly improved fit cannot be simply attributed to the increased number of elasticity parameters associated

with the maximum connectivity. Together with much smaller parameter confidence intervals (Table 4.4, maximum connectivity), we conclude that maximum connectivity is a more realistic assumption for the underlying biochemical pathway than mass-action connectivity.

The approximative capacity of lin-log kinetics is also illustrated by Fig. 4.4. Based on a single optimal reference state, the calculated values matches the measurements closely, bridging large changes in enzyme levels and metabolite concentrations in all 18 different experiments (consider e.g. a 3-fold change of the 2PG concentration between the first and the second set of experiments).

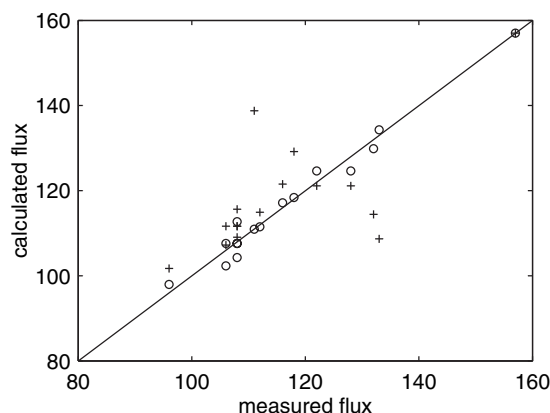
Elasticities describe how a metabolite influences a reaction rate around the reference state; positive elasticities indicate stimulatory effects and negative elasticities indicate inhibitory effects. In the reconstituted metabolic system, the reactions are stimulated by their substrates and inhibited by their products; moreover, BPG stimulates the reaction catalyzed by PGM and ADP stimulates the reaction catalyzed by PK (Giersch, 1995). The signs of the regressed elasticities in Table 4.4 (using the maximum connectivity assumption) comply with this biological knowledge (except the elasticity of PEP for PK, which is very close to zero), without applying any a priori constraints during the linear regression. This demonstrates that the lin-log kinetic formulation is able to capture the essential kinetic information contained in the dataset.

A few more biological interactions are predicted by the estimated elasticities (Table 4.4, maximum connectivity). For example, PEP is predicted to strongly inhibit PGM ($\varepsilon_{\text{PGM,PEP}} = -1.16$). Being a substrate analogue of PGM, the inhibition of the yeast PGM enzyme by PEP has been reported (Sasaki et al., 1971). The inhibition of PK by 2PG has also been described in green alga (Lin et al., 1989). Other interactions are found as well, i.e. BPG stimulates ENO and PK; ADP activates PGM and inhibits ENO. It is important to discuss whether these regressed non-zero elasticities are of biological significance, or merely a result of minimizing the least square errors. To simplify the analysis, the dataset in Table 4.3 are split into two subsets. The first subset contains 16 experiments out of 19 with a constant ADP concentration (i.e. $[\text{ADP}] \approx 107 \mu\text{M}$) so that effects of ADP concentration on the reaction rates are excluded. Similarly, the second subset contains 15 experiments with a constant BPG concentration (i.e. $[\text{BPG}] \approx 13.2 \mu\text{M}$). To find the structural important biological interactions embedded in the two subsets, three different assumptions are tested. Assumption A represents the mass-action assumption, with the independent metabolites BPG and ADP acting specific on the enzymes PGM and PK, respectively. In assumption B these independent metabolites are assumed to be specific as well, but the two intermediates (2PG and PEP) are assumed to affect all three enzymes. Assumption C is identical to the maximum connectivity assumption, where all metabolites are assumed to affect all enzymes. Based on these assumptions, elasticities were estimated from the two sub-datasets as previously described. R_a^2 -values of the best reference state are given in Table 4.5. The fluxes and dependent metabolite concentrations were calculated and compared to the measured ones.

As for subset 1, increased number of parameters has led to better fits, indicated by R_a^2 . In fact, as shown in Fig. 4.5, the calculated fluxes do not deviate from the measured ones only under assumption C. For subset 2 on the other hand, both

Table 4.5 Coefficient of determination with different connectivity assumptions for 2 sub-datasets

| Assumption | Number of elasticities | R_a^2 | |
|------------|------------------------|------------------------------|------------------------------|
| | | Subset 1 ([ADP] constant) | Subset 2 ([BPG] constant) |
| A | 5 | 0.88 | 0.93 |
| B | 7 | 0.97 | 0.99 |
| C | 9 | 0.99 | 0.99 |

**Figure 4.5** Parity plot of measured and calculated fluxes from sub-dataset 1 (in $\mu\text{mol/l/h}$), +: assumption B; o: assumption C.

assumption B and C lead to good agreements between the calculated and measured values of fluxes and dependent metabolite concentrations (not shown). Also, in both subsets the estimated elasticity of PEP on PK is close to zero when assumption B and C are applied (not shown), indicating a negligible effect of PEP on PK. The $K_{m,\text{PEP}}$ of PK is around $50 - 60 \mu\text{M}$ measured at an ADP concentration of 2 mM (Consler et al., 1989; Yu et al., 2003). Since the affinity of PK for PEP increases with decreasing ADP concentration (Ainsworth et al., 1983), the $K_{m,\text{PEP}}$ of PK would be far below $50 \mu\text{M}$ at the ADP concentration used in Giersch's experiments ($60 - 130 \mu\text{M}$). The close to zero elasticity of PEP on PK can therefore be explained by the saturation of PK by PEP at the PEP concentrations used in Giersch's experiments ($50 - 100 \mu\text{M}$).

Summarizing, the above evaluation shows that ADP specifically affects PK while BPG affects not only PGM, but also ENO and PK. Moreover, the elasticity of PEP on PK can be taken as zero. This leads to a reduced number of elasticity parameters (9 compared to 12 in the maximum connectivity assumption), which are required to explain the dataset. These elasticities are then estimated from all 19 experiments as described previously. The results are given in Table 4.4 under 'optimized connectivity' and are used for further calculations. There are no significant changes in the parity plots of the calculated fluxes and dependent metabolite concentra-

tions against the measurements (compared to maximum connectivity assumption in Fig. 4.4) under the optimized connectivity assumption.

Elasticities are enzyme kinetic properties associated with a certain reference state. Table 4.6 lists elasticities at all reference states, which were derived from elasticities at the optimal reference experiment using Eq. 4.14.

FCC estimation using the lin-log approach Table 4.7 gives the FCCs obtained with the lin-log approach. The FCCs given under column A were calculated according to the MCA relationships in Table 4.2, using the elasticities in Table 4.6 for each corresponding reference state. Alternatively, FCCs were directly obtained from the enzyme titration experiment 11-19. While estimated independently using Eq. 4.17, the three FCCs sum up practically to 1 (Table 4.7, column B). Very similar results were obtained when the summation theorem is made implicit with Eq. 4.16 (not shown).

Although the FCCs from the different methods (i.e. indirectly from the elasticities or directly from enzyme titration experiments) differ slightly in values, which should be attributed to the different datasets used (i.e. all 19 experiments vs. experiment 11-19), the distribution of control remains very similar, showing that majority of control is shared by the PK enzyme.

Comparison with results from Giersch Assuming maximum connectivity, Giersch applied enzyme titration experiment 11-19 to estimate elasticities, from which the FCCs were calculated (Giersch, 1995). The same set of data was also used to estimate the FCCs directly, by determining at the reference state the slope of the nonlinear function fitted to enzyme levels and measured fluxes. In Table 4.8, elasticities and FCCs from the lin-log approach and the multiple modulation method by Giersch are compared, at the reference state used by Giersch (i.e. experiment 18).

The elasticity values found by Giersch are somewhat different than those estimated by the lin-log approach. This is not surprising, since the underlying kinetic assumptions (linear vs. linear-logarithmic) and connectivity assumptions (maximum connectivity vs. optimal connectivity) are all different between the multiple modulation and the lin-log approach. It is also worth mentioning that in the maximum connectivity applied by Giersch as such, no assumptions about the specificity of ADP and BPG are necessary (since their concentrations were kept constant in the second set of experiments), while in the optimal connectivity explicit assumptions are made concerning the specificity of these independent metabolites, based on statistical evaluation. Moreover, the data used for parameter estimation are different between the two approaches as well: with multiple modulation, Giersch was only able to use data from the second set of experiments; while with the lin-log approach, all available data have been used as a whole. Despite the difference in values, the sign and magnitude of individual elasticities from the two approaches are in good agreement, the only exception being the elasticities for PK.

The FCCs calculated by the two approaches indirectly from elasticities (Table 4.8, column A) are comparable with each other, showing that PK has the majority of control. C_{PGM}^J from the lin-log approach is slightly negative (-0.17), although experiment 11, 12 and 13 (Table 4.3) indicate that PGM has little or no flux control. The contradiction can be explained by the fact that in the lin-log approach C_{PGM}^J

Table 4.6 Elasticities at different reference states with optimized connectivity assumption^a

| Exp. | PGM | | | | | ENO | | | | | PK | | | | |
|------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|---------------------|---------------------|---------------------|---------------------|--|--|--|
| | $\epsilon_{PGM,2PG}$ | $\epsilon_{PGM,PEP}$ | $\epsilon_{PGM,BPG}$ | $\epsilon_{PGM,ADP}$ | $\epsilon_{ENO,2PG}$ | $\epsilon_{ENO,PEP}$ | $\epsilon_{ENO,BPG}$ | $\epsilon_{ENO,ADP}$ | $\epsilon_{PK,2PG}$ | $\epsilon_{PK,PEP}$ | $\epsilon_{PK,BPG}$ | $\epsilon_{PK,ADP}$ | | | |
| 1 | -0.38 | -0.89 | 1.05 | - | 0.20 | -0.80 | 0.48 | - | -0.21 | - | 0.46 | 0.59 | | | |
| 2 | -0.55 | -1.30 | 1.53 | - | 0.76 | -3.10 | 1.86 | - | -0.25 | - | 0.54 | 0.70 | | | |
| 3 | -0.42 | -0.99 | 1.17 | - | 0.58 | -2.37 | 1.42 | - | -0.19 | - | 0.42 | 0.53 | | | |
| 4 | -0.38 | -0.88 | 1.04 | - | 0.52 | -2.11 | 1.27 | - | -0.17 | - | 0.37 | 0.47 | | | |
| 5 | -0.35 | -0.83 | 0.98 | - | 0.49 | -1.98 | 1.19 | - | -0.16 | - | 0.35 | 0.44 | | | |
| 6 | -0.43 | -1.02 | 1.20 | - | 0.59 | -2.43 | 1.46 | - | -0.19 | - | 0.43 | 0.54 | | | |
| 7 | -0.41 | -0.96 | 1.13 | - | 0.56 | -2.29 | 1.37 | - | -0.18 | - | 0.40 | 0.51 | | | |
| 8 | -0.39 | -0.93 | 1.09 | - | 0.54 | -2.21 | 1.33 | - | -0.18 | - | 0.39 | 0.50 | | | |
| 9 | -0.36 | -0.86 | 1.01 | - | 0.50 | -2.05 | 1.23 | - | -0.16 | - | 0.36 | 0.46 | | | |
| 10 | -0.36 | -0.85 | 1.00 | - | 0.50 | -2.03 | 1.22 | - | -0.16 | - | 0.36 | 0.45 | | | |
| 11 | -0.66 | -1.54 | 1.82 | - | 0.33 | -1.35 | 0.81 | - | -0.23 | - | 0.50 | 0.63 | | | |
| 12 | -0.80 | -1.89 | 2.22 | - | 0.33 | -1.35 | 0.81 | - | -0.23 | - | 0.50 | 0.63 | | | |
| 13 | -0.95 | -2.23 | 2.63 | - | 0.33 | -1.35 | 0.81 | - | -0.23 | - | 0.50 | 0.63 | | | |
| 14 | -0.82 | -1.92 | 2.27 | - | 0.27 | -1.12 | 0.68 | - | -0.23 | - | 0.51 | 0.65 | | | |
| 15 | -0.80 | -1.89 | 2.22 | - | 0.33 | -1.35 | 0.81 | - | -0.23 | - | 0.50 | 0.63 | | | |
| 16 | -0.77 | -1.82 | 2.14 | - | 0.38 | -1.55 | 0.93 | - | -0.22 | - | 0.48 | 0.61 | | | |
| 17 | -0.90 | -2.12 | 2.50 | - | 0.37 | -1.52 | 0.91 | - | -0.23 | - | 0.51 | 0.65 | | | |
| 18 | -0.82 | -1.92 | 2.27 | - | 0.34 | -1.38 | 0.83 | - | -0.23 | - | 0.51 | 0.65 | | | |
| 19 | -0.75 | -1.76 | 2.07 | - | 0.31 | -1.26 | 0.76 | - | -0.23 | - | 0.50 | 0.64 | | | |

^a Elasticities which are set to zero, due to the connectivity assumption made, are indicated by ‘-’

Table 4.7 FCCs estimated by the lin-log approach. A, ‘indirectly’ from elasticities; B, directly from enzyme titration experiment 11-19 with Eq. 4.17

| Exp. nr. | C_{PGM}^J | | C_{ENO}^J | | C_{PK}^J | | Sum of FCCs ^b |
|---------------------|--------------------|-------|--------------------|------|-------------------|------|--------------------------|
| | A | B | A | B | A | B | B |
| 1 | -0.34 | | 0.38 | | 0.96 | | |
| 2 – 10 ^a | -0.34 | | 0.14 | | 1.20 | | |
| 11 | -0.21 | -0.04 | 0.24 | 0.10 | 0.97 | 0.94 | 1.00 |
| 12 | -0.17 | -0.03 | 0.23 | 0.10 | 0.93 | 0.94 | 1.00 |
| 13 | -0.14 | -0.03 | 0.23 | 0.10 | 0.91 | 0.94 | 1.01 |
| 14 | -0.16 | -0.03 | 0.27 | 0.12 | 0.89 | 0.92 | 1.01 |
| 15 | -0.17 | -0.03 | 0.23 | 0.10 | 0.93 | 0.94 | 1.00 |
| 16 | -0.17 | -0.04 | 0.20 | 0.09 | 0.97 | 0.98 | 1.03 |
| 17 | -0.15 | -0.03 | 0.21 | 0.09 | 0.94 | 0.92 | 0.98 |
| 18 | -0.17 | -0.03 | 0.23 | 0.10 | 0.93 | 0.92 | 0.99 |
| 19 | -0.18 | -0.04 | 0.25 | 0.10 | 0.93 | 0.93 | 0.99 |

^a The enzyme levels were the same in experiment 2 to 10, which gives identical FCCs.

^b FCCs estimated ‘indirectly’ (under A) necessarily sum up to 1.

Table 4.8 Elasticities and FCCs estimated by the multiple modulation method and the lin-log approach with experiment 18 as a reference state. The FCCs are obtained: A, ‘indirectly’ from elasticities; B, directly from enzyme titration experiment 11-19.

| Elasticities | Multiple modulation | | Lin-log | |
|---------------------------------------|---------------------|------|---------|----------------|
| $\varepsilon_{\text{PGM},2\text{PG}}$ | -2.83 | | -0.82 | |
| $\varepsilon_{\text{PGM},\text{PEP}}$ | -1.65 | | -1.92 | |
| $\varepsilon_{\text{EBO},2\text{PG}}$ | 0.03 | | 0.34 | |
| $\varepsilon_{\text{ENO},\text{PEP}}$ | -1.35 | | -1.38 | |
| $\varepsilon_{\text{PK},2\text{PG}}$ | 0.01 | | -0.23 | |
| $\varepsilon_{\text{PK},\text{PEP}}$ | 0.29 | | 0 | |
| FCCs | A | B | A | B ^a |
| C_{PGM}^J | 0 | 0 | -0.17 | -0.03 |
| C_{ENO}^J | 0.17 | 0.16 | 0.23 | 0.10 |
| C_{PK}^J | 0.83 | 1.00 | 0.93 | 0.92 |

^a Estimated with Eq. 4.18.

is calculated based on elasticities, which are in turn estimated from a much larger experimental dataset. From the estimated elasticities it can be seen that 2PG inhibits PK. The concentration of 2PG increases with an increasing PGM level (experiment 11-13, Table 4.3), thereby exerting more inhibition on PK. Since PK has the majority of flux control, increase in the PGM level will indirectly lead to a decreased flux. Using $C_{\text{PGM}}^J = -0.17$ and Eq. 4.17, the expected flux change in experiment 11 and 13 is calculated to be less than 4% compared to the reference flux in experiment 12,

showing that the contradiction between a negative C_{PGM}^J and experiment 11-13 is probably only apparent.

The FCCs estimated by the two approaches directly from enzyme titration experiments (Table 4.8, column B) are in good agreement with each other. In the slope-based method however, the summation theorem is often violated when all enzymes in the pathway are titrated (the FCCs determined by Giersch have a sum of more than 1). In the lin-log approach, the summation theorem is much better observed when all FCCs are estimated independently (using Eq. 4.17). Moreover, the lin-log approach allows the summation theorem to be incorporated in the estimation procedure (using Eq. 4.18), which makes the FCCs necessarily sum up to 1 while utilizing all available enzyme titration data.

4.5 Conclusion

The lin-log kinetic approximation has been applied to estimate elasticities and FCCs from a literature dataset. Without any a priori constraints, the lin-log approach is able to deliver biologically sound parameter estimations, although further improvements can be made by implementing biochemical knowledge during the regression. The estimated elasticities are comparable in signs and magnitudes with those obtained by Giersch by the multiple modulation method. The calculated FCCs are very close for both methods. However, the lin-log kinetic approach presented here has several clear advantages compared to the traditional MCA methods to obtain elasticities and control coefficients.

First, the nonlinear lin-log kinetic formulation allows better approximation when large changes in e.g. enzyme levels or independent metabolite concentrations are made. This is evidenced in the good agreement between the calculated and measured fluxes and dependent metabolite concentrations, despite the large perturbations made to the reconstituted pathway (Fig. 4.4). As pointed out by Giersch, the traditional experimental approaches for MCA parameter estimation suffer from difficulties in getting the proper experimental data, which requires small changes to be made in the system and small differences to be measured. Application of the lin-log approach would largely reduce the experimental effort and improve the quality of the estimate, as the influence of experimental scatter would reduce when larger changes are made.

Second, the lin-log approach allows a more straightforward use of experimental data through simple (multiple) linear regression to obtain elasticities and avoids the complicated indirect nonlinear fitting and extrapolation procedures inherent to the traditional MCA approach. Also, it is possible to combine several datasets within one regression based on a single reference state, which gives the best fit. This is not possible with the traditional MCA approach, as it requires that the system variables (in this case the enzyme levels or independent metabolite concentrations) be changed one at a time. Also, the lin-log kinetic approach allows to transform elasticities estimated at one reference state into elasticities and control coefficients for other reference states.

The statistical evaluation of the parameter estimate is not trivial within the traditional MCA framework. Error bounds have been obtained by means of e.g.

Monte-Carlo simulations (Ainscow and Brand, 1998; Ehlde and Zacchi, 1996). Instead, the lin-log approach allows a simple and rigorous statistical evaluation as a result of the linear regression. Different model assumptions can be discriminated on a statistical basis, such as the connectivity assumptions investigated in this study.

In conclusion, the lin-log approach provides a promising alternative to the standard MCA parameter estimation procedures for reliable parameterization of the elasticities, and hence the control coefficients, using large experimental perturbations.

Nomenclature

| | |
|-------------------|---|
| c | concentration of independent metabolite |
| C^J | flux control coefficient |
| C^x | metabolite control coefficient |
| e | enzyme level |
| \mathbf{E}^{c0} | matrix containing elasticities of independent metabolites |
| \mathbf{E}^{x0} | matrix containing elasticities of dependent metabolites |
| J | flux |
| k_{cat} | catalytic constant |
| K_m | affinity in Michaelis-Menten kinetics |
| m | number of dependent metabolites |
| n | number of reactions |
| N | number of observations |
| P | number of parameters |
| a, p, q | parameters in the lin-log representation |
| r | number of independent metabolites |
| R^J | flux response coefficient |
| R^x | metabolite response coefficient |
| R^2 | coefficient of determination |
| R_a^2 | adjusted coefficient of determination |
| \mathbf{S} | stoichiometric matrix |
| SSE | sum of squares due to error |
| SST | sum of squares total |
| v | reaction rate |
| x | concentration of dependent metabolite |

Greek letters

| | |
|-----------------|--------------------------------------|
| ε^c | elasticity of independent metabolite |
| ε^x | elasticity of dependent metabolite |

Superscripts

| | |
|---|---------------------------|
| 0 | chosen reference state |
| , | arbitrary reference state |

References

- Ainscow EK, Brand MD. 1998. Errors associated with metabolic control analysis. Application of Monte-Carlo simulation of experimental data. *J Theor Biol* 194: 223–33.
- Ainsworth S, Kinderlerer J, Gregory RB. 1983. The regulatory properties of rabbit muscle pyruvate kinase. The influence of substrate concentrations. *Biochem J* 209: 401–11.
- Bailey JE. 1991. Toward a science of metabolic engineering. *Science* 252: 1668–75.
- Consler TG, Woodard SH, Lee JC. 1989. Effects of primary sequence differences on the global structure and function of an enzyme: a study of pyruvate kinase isozymes. *Biochemistry* 28: 8756–64.
- Draper NR, Smith H. 1998. *Applied Regression Analysis*. Wiley, New York, 3rd edition.
- Ehlde M, Zacchi G. 1996. Influence of experimental errors on the determination of flux control coefficients from transient metabolite concentrations. *Biochem J* 313: 721–7.
- Eisenhauer JG. 2003. Regression through the origin. *Teaching Statistics* 25: 76–80.
- Fell D. 1997. *Understanding the control of metabolism*. Portland press, London.
- Giersch C. 1995. Determining elasticities from multiple measurements of flux rates and metabolite concentrations. application of the multiple modulation method to a reconstituted pathway. *Eur J Biochem* 227: 194–201.
- Heinrich R, Rapoport TA. 1974. A linear steady-state treatment of enzymatic chains. general properties, control and effector strength. *Eur J Biochem* 42: 89–95.
- Heinrich R, Schuster S. 1996. *The regulation of cellular systems*. Chapman & Hall, New York.
- Hellingwerf KJ, Arents JC, Scholte BJ, Westerhoff HV. 1979. Bacteriorhodopsin in liposomes. II. Experimental evidence in support of a theoretical model. *Biochim Biophys Acta* 547: 561–82.
- Hofer T, Heinrich R. 1993. A second-order approach to metabolic control analysis. *J Theor Biol* 164: 85–102.
- Kacser H, Burns JA. 1973. The control of flux. *Symp Soc Exp Biol* 27: 65–104.
- Lin M, Turpin DH, Plaxton WC. 1989. Pyruvate kinase isozymes from the green alga, *Selenastrum minutum*. II. Kinetic and regulatory properties. *Arch Biochem Biophys* 269: 228–38.
- Nielsen J. 1997. Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates. *Biochem J* 321: 133–8.
- Nielsen J. 1998. Metabolic engineering: techniques for analysis of targets for genetic manipulations. *Biotechnol Bioeng* 58: 125–32.
- Onsager L. 1931. Reciprocal relations in irreversible processes. *Phys Rev* 37: 405–26.
- Petterson G. 1996. Error associated with experimental flux control coefficient determinations in the Calvin cycle. *Biochim Biophys Acta* 1289: 169–74.
- Rottenberg H. 1973. The thermodynamic description of enzyme-catalyzed reactions. The linear relation between the reaction rate and the affinity. *Biophys J* 13: 503–11.
- Sasaki K, Sugimoto E, Chiba H. 1971. Studies on the active site of yeast phosphoglycerate mutase. *Biochim Biophys Acta* 227: 584–94.
- Simpson T, Shimizu H, Stephanopoulos G. 1998. Experimental determination of group flux

- control coefficients in metabolic networks. *Biotechnol Bioeng* 58: 149–53.
- Small JR. 1993. Flux control coefficients determined by inhibitor titration: the design and analysis of experiments to minimize errors. *Biochem J* 296: 423–33.
- Small JR, Kacser H. 1993a. Responses of metabolic systems to large changes in enzyme activities and effectors. 1. The linear treatment of unbranched chains. *Eur J Biochem* 213: 613–24.
- Small JR, Kacser H. 1993b. Responses of metabolic systems to large changes in enzyme activities and effectors. 2. The linear treatment of branched pathways and metabolite concentrations. Assessment of the general non-linear case. *Eur J Biochem* 213: 625–40.
- van Dam K, Westerhoff HV, Rutgers M, Bode JA, de Jonge PC, Bos MM, van den Berg G. 1981. Comparison of reaction stoichiometry to differential importance of thermodynamic forces for reaction rates. In: Palmieri F, Quagliariello E, Siliprandi N, Slater EC (eds.), *Vectorial reactions in electron and ion transport in mitochondria and bacteria*. Elsevier.
- van der Meer R, Westerhoff HV, Van Dam K. 1980. Linear relation between rate and thermodynamic force in enzyme-catalyzed reactions. *Biochim Biophys Acta* 591: 488–93.
- Visser D, Heijnen JJ. 2002. The mathematics of metabolic control analysis revisited. *Metab Eng* 4: 114–23.
- Visser D, Heijnen JJ. 2003. Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metab Eng* 5: 164–76.
- Westerhoff HV, Hellingwerf KJ, Arents JC, Scholte BJ, Van Dam K. 1981. Mosaic nonequilibrium thermodynamics describes biological energy transduction. *Proc Natl Acad Sci U S A* 78: 3554–8.
- Westerhoff HV, Lolkema J, Otto R, Hellingwerf KJ. 1982. Thermodynamics of growth. Non-equilibrium thermodynamics of bacterial growth. The phenomenological and the mosaic approach. *Biochim Biophys Acta* 683: 181–220.
- Westerhoff HV, van Dam K. 1987. *Thermodynamics and control of biological free energy transduction*. Elsevier, Amsterdam.
- Wiechert W. 2002. Modeling and simulation: tools for metabolic engineering. *J Biotechnol* 94: 37–63.
- Yu S, Lee LL, Lee JC. 2003. Effects of metabolites on the structural dynamics of rabbit muscle pyruvate kinase. *Biophys Chem* 103: 1–11.

Appendix A

Consider the well-known Michaelis-Menten kinetics:

$$v = k_{cat}e \frac{x}{K_m + x} \tag{A1}$$

For the elasticity ε_x one obtains according to Eq. 4.2:

$$\varepsilon_x = \frac{1}{1 + x/K_m} \tag{A2}$$

For the reference state with J^0 , e^0 and x^0 Eq. A2 turn into:

$$J^0 = k_{cat} e^0 \frac{x^0}{K_m + x^0} \quad (\text{A3})$$

$$\varepsilon_x^0 = \frac{1}{1 + x^0/K_m} \quad (\text{A4})$$

Using Eq. A3 and A4, K_m and k_{cat} can be eliminated from Eq. A1, which gives:

$$\frac{v}{J^0} = \frac{e}{e^0} \left(\frac{x/x^0}{\varepsilon_x^0 + (1 - \varepsilon_x^0)x/x^0} \right) \quad (\text{A5})$$

In Fig. 4.1, the linear approximation of Eq. A5:

$$\frac{v}{J^0} - 1 = \left(\frac{e}{e^0} - 1 \right) + \varepsilon_x^0 \left(\frac{x}{x^0} - 1 \right)$$

and the lin-log approximation of Eq. A5:

$$\frac{v}{J^0} = \frac{e}{e^0} \left[1 + \varepsilon_x^0 \ln \left(\frac{x}{x^0} \right) \right]$$

are compared taking $e/e^0 = 1$, $\varepsilon_x^0 = 0.5$.

Appendix B

Eq. 4.15 used for the estimation of elasticities corresponds to the following equation of multiple linear regression:

$$y_i = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4$$

Here y_i represents the left-hand side term of Eq. 4.15, z_i 's are the logarithmic concentrations and β_i ($i \neq 0$) the elasticity parameters. β_0 in the above equation is zero, which results in a special case of linear regression, i.e. regression through the origin (RTO). It holds for RTO that (Eisenhauer, 2003):

$$\Sigma(y_i)^2 = \Sigma(y_i - \hat{y}_i)^2 + \Sigma(\hat{y}_i)^2$$

where \hat{y}_i is the calculated value of y_i from the estimated elasticities, $\Sigma(y_i)^2$ is the sum of squares total (*SST*) and $\Sigma(y_i - \hat{y}_i)^2$ is the sum of squares due to error (*SSE*). The coefficient of determination R^2 can therefore be defined for RTO as

$$R^2 = 1 - \frac{SSE}{SST}$$

Obviously, R^2 equals 1 when the model would perfectly fit the data (i.e. $SSE = 0$). Therefore R^2 can be seen as the percentage of variations in the experimental data explained by regression. In the case of RTO, the degree of freedom of *SST* equals the total number of observations (N). The degree of freedom of *SSE*, on the other hand, also depends on the number of parameters (P) to be estimated and equals $N - P$. When models with different number of parameters are compared, the different

degrees of freedom of SSE can be taken into account during the evaluation of R^2 by defining an adjusted coefficient of determination R_a^2 :

$$R_a^2 = 1 - \frac{SSE/(N - P)}{SST/N} = 1 - (1 - R^2) \frac{N}{N - P}$$

which still equals 1 when a perfect fit is obtained.

Chapter 5

A Conceptual Functional Genomics Strategy Using Solely Metabolome Data

Abstract A gene with yet unknown physiological function can be studied by changing its expression level followed by analysis of the resulting phenotype. This type of functional genomics study can be complicated by the occurrence of ‘silent mutations’, the phenotypes of which are not easily observable in terms of metabolic fluxes (e.g. the growth rate). Nevertheless, genetic alteration may give rise to significant yet complicated changes in the metabolome. We propose here a conceptual functional genomics strategy based on microbial metabolome data, which identifies changes in *in vivo* enzyme activities in the mutants. These predicted changes are used to formulate hypotheses to infer unknown gene functions. The required metabolome data can be obtained solely from high-throughput mass spectrometry analysis, which provides the following *in vivo* information: 1) the metabolite concentrations in the reference and the mutant strain; 2) the metabolic fluxes in both strains and 3) the enzyme kinetic parameters of the reference strain. We demonstrate *in silico* that changes in enzyme activities can be accurately predicted by this approach, even in ‘silent mutants’.

5.1 Introduction

The increasing availability of whole-genome data is paralleled by an ever increasing number of genes with unknown function, urging both the development and integration of functional genomics strategies . Gene functions can be inferred from a variety of data sources, obtained at the level of genome, transcriptome, proteome and metabolome (Castrillo and Oliver, 2004).

Published as: Wu L, van Winden WA, van Gulik WM, Heijnen JJ. 2005. Application of metabolome data in functional genomics: A conceptual strategy. *Metab Eng* 7: 302-10.

Among these ‘omics’-approaches, metabolomics aims at the comprehensive analysis of all metabolites (Fiehn, 2002). The metabolome possesses the advantages of being closely related to phenotypes and being more context-dependent. Furthermore, metabolome analysis can be easily extended to different organisms, since a metabolite has the same chemical structure across all species (Goodacre et al., 2004). The application of metabolome data in functional genomics was initiated by the FANCY approach (Raamsdonk et al., 2001), based on the idea that genetic alterations often lead to significant changes in metabolite concentrations, even in ‘silent mutations’ with no clear flux-phenotypes (e.g. the growth rate). FANCY suggests gene functions by comparing metabolome profiles of strains deleted for known and unknown genes. When the genes with known and unknown function are located in the same monofunctional unit (Rohwer et al., 1996), deletion will result in the same response, e.g. in the metabolite concentrations, in the rest of the metabolic network. The co-response coefficient, defined in the framework of metabolic control analysis (Hofmeyr and Cornish-Bowden, 1996), is used to quantify these responses. To calculate the co-response coefficient with measured metabolite concentrations, where large changes are likely to occur due to gene deletion, a linear approximation is used (Raamsdonk et al., 2001). Genes functions can be inferred based on the calculated approximation of co-response coefficients. When applied to metabolome-wide data, FANCY employs multivariate analysis to infer the similarity of metabolite responses. Likewise, metabolite profiling combined with multivariate statistics has been successfully applied to genotype differentiation in plants and fungi (Allen et al., 2003; Fiehn et al., 2000; Smedsgaard and Nielsen, 2005). Next to multivariate data analysis, Foster et al. (2002) proposed a framework to identify gene functions in single deletion mutants by combining metabolome data with *in silico* pathway analysis.

One of the major disadvantages of applying metabolome data in functional genomics is that multiple changes in the metabolome are only indirectly related to genetic alterations. Instead, they arise directly due to changes in enzyme activity (i.e. the *in vivo* maximum catalytic capacity of an enzyme) and kinetic enzyme-metabolite interactions, the latter often extremely complex. Hence, metabolome data as such do not provide a straightforward insight in gene function, which necessitates data-driven approaches, such as FANCY. To enable an abductive and more focused analysis which complements FANCY, it would be desirable to access enzyme activity changes, which are indicative of either direct or indirect, regulatory effects of the altered gene and bear a closer link to its function. Quantitative proteome analysis has indeed been an integral part of functional genomics investigations in perturbed metabolic networks (Baliga et al., 2002; Dainese et al., 1997; Haurie et al., 2001; Ideker et al., 2001; Knowles et al., 2003). However, the applied techniques focus primarily on protein abundance instead of enzyme activities *in vivo*, hence they overlook possible intracellular regulatory mechanisms (Kobe and Kemp, 1999).

In this paper we develop a conceptual functional genomics strategy, which uses metabolome data only. Unlike the FANCY approach that compares the metabolome data of different strains and suggests unknown gene functions based on monofunctional units and co-response of metabolites, our approach utilizes metabolome data to calculate the relative changes in *in vivo* enzyme activities in the mutant relative to a reference (e.g. a wild type), thereby making use of a quantitative kinetic model. These predicted changes reveal the site(s) of action of the altered gene and are es-

sential to generate sound hypotheses regarding its function, which can be delineated with further experimental investigation. The merit of this strategy is demonstrated with an *in silico* case study.

5.2 Concept of the approach

To calculate *in vivo* enzyme activity changes, we exploit enzyme kinetics that defines the flux (v_i) through an enzyme i as a nonlinear function of enzyme activity (e_i) and metabolite concentrations (\mathbf{x}):

$$v_i = f_i(e_i, \mathbf{x}, \mathbf{P}) \quad (5.1)$$

The *in vivo* enzyme activity can be thus calculated from Eq. 5.1 when metabolite concentrations, fluxes and a set of kinetic parameters \mathbf{P} are obtained *in vivo*, in a suitable kinetic format given by f_i .

The kinetic format can be mechanistic enzyme rate laws, i.e. based on detailed knowledge of catalytic mechanisms, such as the well-known Michaelis-Menten kinetics. In addition, a number of non-mechanistic, approximative kinetic formats have been developed, such as the linear approximation (which has been the basis of Metabolic Control Analysis, see e.g. Delgado and Liao 1992; Heinrich and Rapoport 1977; Kacser and Burns 1973), the log-linear approximation (Hatzimanikatis and Bailey, 1996), the powerlaw approximation (Savageau, 1976; Voit, 2000), the thermokinetic approach (Nielsen, 1997; Westerhoff and van Dam, 1987) and the recently developed linear-logarithmic approximation (Heijnen et al., 2004; Visser and Heijnen, 2003; Wu et al., 2004). The approximative kinetic formats have the advantage of containing much fewer parameters, a uniform structure and universal applicability, also when the catalytic mechanism is unknown. These properties reduce the efforts for model building and parameter estimation. Another desirable property for a suitable kinetic format is good approximative capacity over a large range of metabolite concentrations and fluxes. Hence, nonlinear kinetic formats are preferred above (semi-) linear kinetics defined close to a reference point.

Here we employ the linear-logarithmic (lin-log) kinetics to illustrate our conceptual strategy. Both *in silico* and experimental studies have shown a satisfactory approximative performance of this nonlinear kinetic format upon large changes in metabolite concentrations, enzyme activities and fluxes (Visser and Heijnen, 2003). The lin-log kinetics is given by:

$$\frac{\mathbf{v}}{\mathbf{v}^0} = \left[\frac{\mathbf{e}}{\mathbf{e}^0} \right] \left(\mathbf{i} + \mathbf{E}^{\mathbf{x}^0} \ln \frac{\mathbf{x}}{\mathbf{x}^0} \right) \quad (5.2)$$

where the metabolic fluxes, enzyme activities and metabolite concentrations are expressed relative to a reference state (e.g. the wild type strain, denoted by 0). Relative changes in metabolic fluxes (\mathbf{v}/\mathbf{v}^0) are considered proportional to relative changes in the activities of the corresponding enzymes (\mathbf{e}/\mathbf{e}^0). The nonlinear effects of metabolite concentrations on reaction rates are incorporated in lin-log kinetics as the sum of logarithms of relative changes in metabolite concentrations (\mathbf{x}/\mathbf{x}^0). The matrix $\mathbf{E}^{\mathbf{x}^0}$ contains elasticities, which serve as the kinetic parameters. They are

defined as the scaled sensitivities of reaction rates towards metabolite concentrations at the reference state:

$$\varepsilon_{ik}^{x^0} = \frac{x_k^0}{v_i^0} \left(\frac{\partial v_i}{\partial x_k} \right)^0 \quad (5.3)$$

Thus, one single elasticity parameter per metabolite is involved in each reaction, thereby largely reduce the number of parameters to be identified. Calculation of relative changes in *in vivo* enzyme activities (\mathbf{e}/\mathbf{e}^0) from Eq. 5.2 therefore requires: 1) the metabolite concentrations in the mutant (\mathbf{x}) and reference state (\mathbf{x}^0), 2) the metabolic fluxes in the mutant (\mathbf{v}) and reference state (\mathbf{v}^0) and 3) the elasticities (\mathbf{E}^{x^0}) in the reference state only.

All the required information can be obtained solely from mass spectrometry (MS) analysis of the metabolome. To obtain *in vivo* metabolite concentrations (\mathbf{x} and \mathbf{x}^0), cells are rapidly sampled (Lange et al., 2001; Schaefer et al., 1999; Theobald et al., 1993) and immediately quenched to prevent further turnover of the metabolites (de Koning and van Dam, 1992; Mashego et al., 2003). Cells are subsequently disrupted to release intracellular metabolites (de Koning and van Dam, 1992; Gonzalez et al., 1997; Theobald et al., 1997). The obtained cell-free extracts containing intracellular metabolites can be analyzed by e.g. LC-MS (Buchholz et al., 2001; van Dam et al., 2002). Compared to genome-wide transcriptome analysis, spatial and temporal profiling of the metabolome by MS is still in an early stage of development. However, promising advances have been made to improve the precision of LC-ESI-MS analysis by incorporating metabolome-wide fully U-¹³C-labeled internal standards (Wu et al., 2005) and portions of the total metabolome can now be identified and quantified (Goodacre et al., 2004; Soga et al., 2002).

To obtain *in vivo* (quasi) steady state fluxes (\mathbf{v}^0) in a metabolic sub-network with known stoichiometry, isotopic tracer experiments (e.g. with ¹³C-labeled molecules) are required (Wiechert, 2001). Cells are fed with labeled metabolites to reach an isotopic (quasi-) stationary state. The mass isotopomer distributions of the ¹³C-metabolome, e.g. the proteinogenic amino acids (Fischer and Sauer, 2003; Gombert et al., 2001), excreted metabolites (Wittmann and Heinzle, 2001) or free intracellular metabolites (Kromer et al., 2004; van Winden et al., 2005), are determined by MS. These data are subsequently applied to estimate *in vivo* fluxes using well-established computational tools (Wiechert et al., 2001).

To obtain the kinetic parameters (\mathbf{E}^{x^0}), stimulus response experiments are required. In such experiments, the metabolism of intact cells in a reference (quasi-) steady state is perturbed by external stimuli (e.g. a substrate pulse) (Chassagnole et al., 2002; Rizzi et al., 1997; Vaseghi et al., 1999; Visser et al., 2004). The subsequent metabolome responses are followed in a short time window, employing the experimental procedures described above to obtain *in vivo* metabolite concentrations. By limiting the observation time window, it can be assumed that the metabolome responses are primarily triggered by kinetic interactions and that enzyme activities are essentially constant (i.e. $\mathbf{e}(t)/\mathbf{e}^0 \approx \mathbf{i}$) (Theobald et al., 1997). The measured temporal evolution of metabolite concentrations ($\mathbf{x}(t)$) is combined with the measured \mathbf{x}^0 to estimate the elasticities with available mathematical procedures (Haunschild et al., 2005; Kresnowati et al., 2005; Rizzi et al., 1997).

Summarizing, an MS-based metabolomics platform in principle suffices to provide the required information to calculate \mathbf{e}/\mathbf{e}^0 . Unlike the well-established and

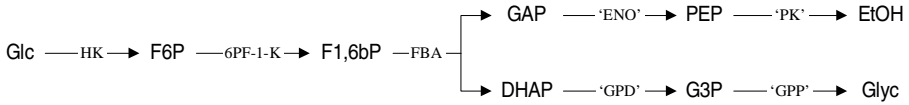


Figure 5.1 **Metabolic network of the *in silico* strains.** ‘HK’ represents the lumped reactions catalyzed by HK and PGI; ‘ENO’ represents the lumped reactions catalyzed by GAPDH, PGK, PGM and ENO; ‘PK’ represents the lumped reactions catalyzed by PK, pyruvate decarboxylase and alcohol dehydrogenase. ‘GPD’ represents the two glycerol 3-phosphate dehydrogenase isoenzymes encoded by *GPD1* and *GPD2*; ‘GPP’ represents the two glycerol-3-phosphatase isoenzymes encoded by *GPP1* and *GPP2*.

widely applied ^{13}C flux analysis, the unbiased quantification of a large number of intracellular metabolites and the construction of kinetic models are more challenging (Wiechert, 2002). The latter not only requires high-throughput metabolome analysis, but also a priori knowledge of existing *in vivo* (allosteric) interactions between enzymes and metabolites. Despite decades of *in vitro* biochemical research, this information is yet far from complete. Hence, the development and validation of *in vivo* kinetic models of metabolic sub-networks of interest is crucial to the practical implementation of the proposed approach.

5.3 The *in silico* model

We performed an *in silico* case study to test whether changes in enzyme activities can be correctly calculated, when the required information is available for a reference wild type strain and silent or not-silent mutants. The *in silico* ‘wild type’ (in a *S. cerevisiae* background) was adapted from Teusink et al. (1998) and consisted of 7 (lumped) enzymatic steps in glycolysis converting glucose to glycerol and ethanol (Fig. 5.1). Each step was assigned a mechanistic rate equation, representing the ‘true’ enzyme kinetics (see Appendix).

Three different genes of *S. cerevisiae* were chosen and assumed here to have as yet unknown functions. Hence, we consider three different single-gene mutants to illustrate our approach:

- 1) The *PFK27* gene encoding 6-phosphofructo-2-kinase (6PF-2-K) (Kretschmer and Fraenkel, 1991) was assumed to have yet unknown function. 6PF-2-K converts fructose-6-phosphate (F6P) into fructose-2,6-bisphosphate (F2,6bP), an activator of 6-phosphofructo-1-kinase (6PF-1-K) (Nissler et al., 1983). In the *pfk27+* mutant, *PFK27* is overexpressed, leading to a 50% increase of the F2,6bP concentration.
- 2) The *GPD1* gene encoding an isoenzyme of glycerol phosphate dehydrogenase (GPD) (Albertyn et al., 1994) was assumed to have yet unknown function. In its deletion mutant ($\Delta gpd1$), the abundance of GPD is at 30% of its wild type level. The remaining GPD activity is provided by the isoenzyme Gpd2p (Eriksson et al., 1995). The responses of the first two mutants have been previously described in (Teusink et al., 1998).
- 3) The *X* gene encoding a regulator of the *GPD1* and *GPP2* gene expression.

The *GPP1* and *GPP2* genes encode two glycerol-3-phosphatase isoenzymes (Pahlman et al., 2001). In its deletion mutant (Δx), the abundance of Gpd1p and Gpp2p is affected such that the total activities of ‘GPD’ and ‘GPP’ are reduced by 50% each.

The chosen mutants reflect different mechanisms by which enzyme activities are affected by gene alterations, either direct (in the $\Delta gpd1$ mutant) or indirect, e.g. through the effect of an allosteric effector (in the mutant *pfk27+*) or gene regulatory mechanisms (in the mutant Δx). Moreover, these gene alterations lead to different flux phenotypes, silent or non-silent, as will be shown later. The chosen set of mutant therefore provides a suitable yet simple benchmark for the approach to be tested. For each mutant, the proposed metabolomics-based approach should thus be able to distinguish those few enzymes that are affected by the genetic changes from the majority of non-affected ones. Further it should deliver a qualitatively correct prediction of the enzyme activity changes, such that useful hypotheses can be generated regarding the unknown gene function.

5.4 Methods

Steady state metabolite concentrations and fluxes of the *in silico* wild type and mutants were obtained by implementing and simulating the following mass balances in Matlab (The Mathworks Inc., Natick, MA, USA):

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v} \tag{5.4}$$

\mathbf{S} is a stoichiometric matrix obtained from the metabolic network depicted in Fig. 5.1. \mathbf{v} is a vector containing all reaction rates as specified by the mechanistic rate equations. The concentration of glucose, ethanol and glycerol were fixed at 20 mM, 0 mM and 0 mM respectively. An initial concentration of 1 mM was used for all other metabolites in the simulation. The F2,6bP concentration in the wild type, $\Delta gpd1$ and Δx strains was set to 1 mM, in the *pfk27+* mutant it was increased by 50% to 1.5 mM. It was further assumed that the flux through 6PF-2-K is negligible.

The elasticities in the lin-log kinetic equations were derived from the mechanistic rate equations according to Eq. 5.3, taking the steady state of the wild type as a reference state. For 6PF-1-K, the lin-log rate equation did not include F2,6bP as an effector, as *PFK27* was treated essentially as a gene with unknown function. Relative changes of enzyme activities were subsequently solved from Eq. 5.2 for all three mutants.

5.5 Results and discussion

The metabolite concentrations and fluxes in the wild type and the mutant strains, obtained from simulation, are given in Table 5.1. Nearly all metabolite concentrations have changed in the three mutant strains, up to a factor of 4 compared to the wild type, due to the change of only one or two enzymes. This illustrates the necessity of the proposed approach, which seeks to deduce from the otherwise complex and unrevealing metabolome data a limited set of enzymes with changed activity. A

Table 5.1 Simulated steady state metabolite concentrations (mM) and flux ($\mu\text{mol}/1/\text{min}$) for the *in silico* wild type and mutant strains

| Metabolites | WT | <i>pfk27+</i> | Δgpd1 | Δx |
|-------------------|------|---------------|---------------------|------------|
| F6P | 3.22 | 1.74 | 3.22 | 3.22 |
| F1,6bP | 3.07 | 3.82 | 6.55 | 5.56 |
| GAP | 0.53 | 0.51 | 0.30 | 0.32 |
| PEP | 0.28 | 0.23 | 0.10 | 0.12 |
| DHAP | 0.37 | 0.46 | 1.30 | 1.06 |
| G3P | 0.16 | 0.19 | 0.16 | 0.38 |
| v_{HK}^a | 0.69 | 0.80 | 0.69 | 0.69 |

^a The flux through ‘HK’, which is identical to the fluxes through 6PF-1-K and FBA and two times the fluxes through ‘GPD’, ‘GPP’, ‘ENO’ and ‘PK’, due to the stoichiometric constraint.

Table 5.2 Calculated elasticities for the *in silico* wild type strain as a reference state

| E^{x0} | Glc | F6P | F1,6bP | GAP | PEP | DHAP | G3P |
|----------|------|-------|--------|-------|-------|-------|-------|
| ‘HK’ | 0.37 | -0.32 | 0 | 0 | 0 | 0 | 0 |
| 6PF-1-K | 0 | 0.79 | 0 | 0 | 0 | 0 | 0 |
| FBA | 0 | 0 | 2.10 | -2.13 | 0 | -2.05 | 0 |
| ‘ENO’ | 0 | 0 | 0 | 1.82 | -1.27 | 0 | 0 |
| ‘PK’ | 0 | 0 | 2.06 | 0 | 1.38 | 0 | 0 |
| ‘GPD’ | 0 | 0 | 0 | 0 | 0 | 1.51 | -0.86 |
| ‘GPP’ | 0 | 0 | 0 | 0 | 0 | 0 | 0.86 |

flux-phenotype is only observed for *pfk27+*, due to the increased level of F2,6bP, which activates 6PF-1-K. The Δgpd1 and Δx mutants are ‘silent’ with respect to the fluxes, due to the lack of triose phosphate isomerase, which ensures a fixed 1:1 flux ratio towards glycerol and ethanol set by FBA, and the absence of feedback effects of metabolites downstream of F1,6bP on ‘HK’ and 6PF-1-K, which ensures a constant glucose uptake rate.

Elasticities calculated from the kinetic model, with the wild type as a reference state, are given in Table 5.2. The lin-log kinetic model contains a mere 12 elasticity parameters, a reduction of about 50% compared to mechanistic rate equations (excluding the v_f ’s as they are normalized in the lin-log kinetics). The modulatory effect of each metabolite on the reaction rate can be immediately seen from the sign of its elasticity, i.e. positive elasticities indicate stimulatory effects (as can be found for substrates and activators of the reaction), while negative elasticities indicate inhibitory effects (as can be found for products of the reaction).

Fig. 5.2 shows the calculated relative enzyme activity changes in the mutants compared to the wild type strain. In all *in silico* mutants, the enzymes affected by genetic perturbation can be clearly distinguished from the non-affected ones, as the activity changes of the latter are all calculated to be relatively close to 1. For the

pfk27+ mutant, a nearly 2-fold increase in 6PF-1-K activity is found, which corresponds to the stimulatory effect of the increased concentration of F2,6bP. For the two ‘silent mutants’, the calculated e/e^0 of the affected enzymes are quantitatively correct: in the $\Delta gpd1$ mutant, the GPD activity is calculated to be 35% of the wild type level, compared to a 30% implemented in the simulation; in the Δx mutant, the GPD and GPP activities are calculated to be 54% and 57% of their wild type level respectively, compared to a 50% implemented for both enzyme activities in the simulation. Thus, the proposed method using approximative lin-log kinetics appears robust and its predictive capacity is satisfactory for *in silico* strains tested, including the two ‘silent mutants’.

These calculated changes in enzyme activities allow a more focused analysis of the function of the altered gene, by means of formulation of hypotheses and the subsequent delineation of them. This is a process highly dependent on background biochemical knowledge of the metabolic sub-network. For the *in silico* strains tested, three relevant mechanisms can be recognized, leading to calculated changes in enzyme activity: A. changes in enzyme abundance; B. changes in post-translational modification, e.g. the enzyme is (de)phosphorylated; C. concentration changes of yet unknown effector metabolite(s). All three mechanisms naturally extend to other enzymes having the same catalytic activity as the affected ones. The altered gene can either directly cause or indirectly induce these changes in enzyme activities, leading to a combinatorial set of hypotheses regarding its function.

Using *in vitro* enzyme assays only, initial delineation of hypotheses can lead to well-based conclusions, which form leads for further experimental investigation. For example, in the *pfk27+* mutant, a 6PF-1-K assay should reveal an unchanged enzyme activity, suggesting unchanged abundance and phosphorylation status of 6PF-1-K, which contradicts the calculated increase in e/e^0 . To simplify matter, we exclude here the possibility of a combined but opposite change of enzyme abundance and phosphorylation, which can be experimentally verified using either 2D-gel electrophoresis or isotope-coded affinity tagging (Gygi et al., 1999). This rules out mechanisms A and B. Hence the *PFK27* gene is most likely associated with ‘yet unknown’ effector(s) of 6PF-1-K (i.e. mechanism C), which might have gone unnoticed if *only* the enzyme abundance were quantified. For the $\Delta gpd1$ and the Δx mutant, enzyme assays should confirm the calculated enzyme activity changes. For Gpd2p, Gpp1p and Gpp2p (Gpd1p is yet an ‘unknown’ gene product), there have been yet no post-translational processes reported in *S. cerevisiae*. Thus, most likely, in the $\Delta gpd1$ mutant the deleted gene either encodes or induces transcriptional regulation (in this case repression) of an enzyme with GPD activity, assuming a correct kinetic model. The former hypothesis is supported by the otherwise inconclusive sequence comparison, which shows that *GPD2* is 69% identical to the ‘unknown’ gene (Eriksson et al., 1995). Likewise, the enzyme activity changes in the Δx mutant are indicative of changes in enzyme abundance. As both affected enzymes participate in glycerol biosynthesis, the tentative hypothesis could be that they are co-regulated by gene *X*.

In Table 5.3 our approach is compared to the FANCY approach in terms of required information input and deliverables. Both approaches rely on high-throughput metabolome analysis. The FANCY approach requires a large library of metabolome profiles of mutants deleted for known genes. In our approach a well-characterized

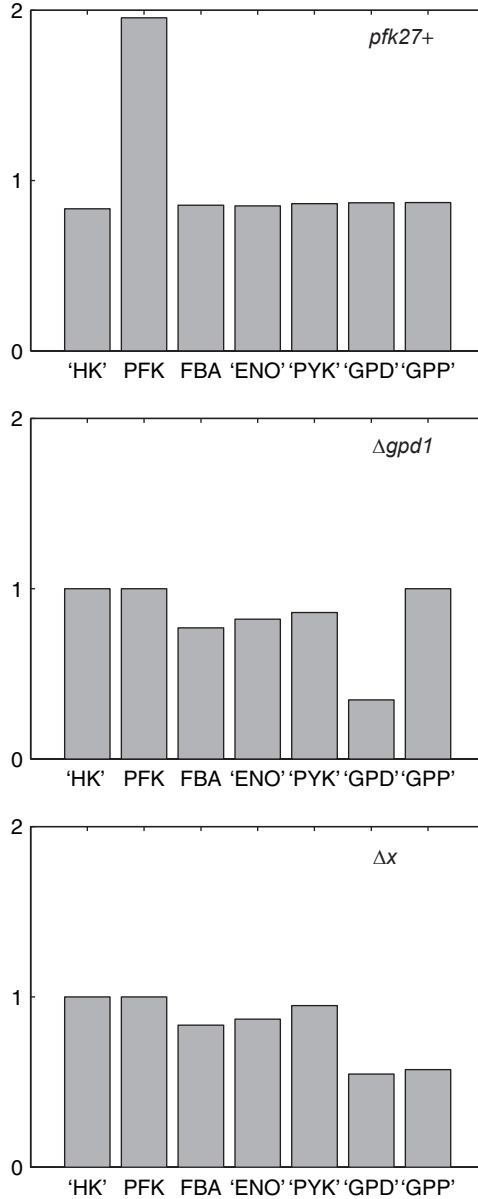


Figure 5.2 Calculated relative changes in enzyme activities (e/e^0) in the mutant strains compared to the wild type.

wild type strain, in terms of metabolite concentrations, fluxes and *in vivo* kinetics, serves as a reference state that can be used in principle against any mutants belonging to the same species. For the mutant of interest, the *in vivo* fluxes should be known next to its metabolite concentrations. Central to our approach is a non-linear approximative kinetic model of a metabolic sub-network, which replaces the assumption on monofunctional units and the linear approximation of co-response

Table 5.3 Comparison of the information input and output between the FANCY approach and current approach

| Approach | | FANCY | | Current approach | |
|--------------|----------------|--------------------------------------|--|---|--|
| Strains used | | Mutants of genes with known function | Mutant of a gene with unknown function | Wild type | Mutant of a gene with unknown function |
| Input | \mathbf{x} | + | + | + | + |
| | \mathbf{v}^a | – | – | + | + |
| | f_i^a | – | – | + | – |
| Output | | Suggested similarity to known genes | | Calculated changes in enzyme activities | |

^a Metabolome derived *in vivo* fluxes (\mathbf{v}) and kinetic model (f_i).

coefficients in the FANCY approach.

In the simple *in silico* case studies presented, a limited number of enzymes is considered to have activity changes. Both the Δx mutant and the $\Delta gpd1$ mutant contain changes due to a ‘direct’ effect of the gene alteration (Δx can be considered to have a direct ‘regulatory’ effect), while the *pfk27+* mutant contain both a direct effect (i.e. increase of the Pfk27p activity and the concentration of F2,6bP) and an indirect effect (i.e. F2,6bP activation of PFK). In practice however, a single gene alteration might bring about many yet unknown, indirect effects in an integrated metabolic network and simultaneously affect the activity of various enzymes, which can complicate the analysis of the unknown gene function.

Targets might be selected for in-depth analytical scrutiny, either based on the magnitude of their calculated activity changes or the completeness of biochemical information of affected enzymes. In addition, the number of unknown indirect effects can be limited by choosing a well-studied ‘core’ metabolic sub-network to generate working hypotheses. An example is the primary metabolism of *S. cerevisiae*, for which abundant literature on kinetic interactions and regulatory mechanisms is present. Moreover, genetic perturbations are likely to be reflectible on enzyme activity changes in the primary metabolism, due to its tight regulation and high connectivity with other modules of the metabolic network. Understanding of a well-characterized sub-network can be further synergetic with functional genomics studies, as the latter allows a continuous extension and iterative improvement of sub-network models (e.g. in the case of the *pfk27+* and $\Delta gpd1$ mutant). The unknown indirect effects can be further limited by combining transcriptome analysis of the mutant and wild type with the predicted enzyme activities changes, although here cautions must be taken that enzyme activities and expression levels are not necessarily correlated (Ideker et al., 2001). For example, a predicted increase of enzyme activity might indicate an indirect effect if this corresponds to an increased expression of the encoding gene or a positive regulator of it. Finally, it has been shown that hypothesis generation and delineation can be made more efficient by implementing well-programmed artificial intelligence algorithms (King et al., 2004).

In conclusion, we have demonstrated *in silico* that the proposed strategy is able

to identify changes in enzyme activities upon changes in genes that are assumed to have yet unknown function, using solely metabolome derived information in a reference wild type strain and the mutant. The practical implementation of the proposed strategy will largely rely on high-throughput and unbiased metabolome quantification and the availability of validated *in vivo* kinetic models. We expect that integration of the predicted enzyme activity changes with data emerging from other individual 'omics'-approaches will accelerate the elucidation of unknown gene functions in a model-based, hypothesis-driven fashion and complement the existing, data-driven FANCY approach.

References

- Albertyn J, Hohmann S, Thevelein JM, Prior BA. 1994. GPD1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in *Saccharomyces cerevisiae*, and its expression is regulated by the high-osmolarity glycerol response pathway. *Mol Cell Biol* 14: 4135–44.
- Allen J, Davey HM, Broadhurst D, Heald JK, Rowland JJ, Oliver SG, Kell DB. 2003. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat Biotechnol* 21: 692–6.
- Baliga NS, Pan M, Goo YA, Yi EC, Goodlett DR, Dimitrov K, Shannon P, Aebersold R, Ng WV, Hood L. 2002. Coordinate regulation of energy transduction modules in *Halobacterium sp.* analyzed by a global systems approach. *Proc Natl Acad Sci U S A* 99: 14913–8.
- Buchholz A, Takors R, Wandrey C. 2001. Quantification of intracellular metabolites in *Escherichia coli* K12 using liquid chromatographic-electrospray ionization tandem mass spectrometric techniques. *Anal Biochem* 295: 129–37.
- Castrillo JI, Oliver SG. 2004. Yeast as a touchstone in post-genomic research: strategies for integrative analysis in functional genomics. *J Biochem Mol Biol* 37: 93–106.
- Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M. 2002. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol Bioeng* 79: 53–73.
- Dainese P, Staudenmann W, Quadroni M, Korostensky C, Gonnet G, Kertesz M, James P. 1997. Probing protein function using a combination of gene knockout and proteome analysis by mass spectrometry. *Electrophoresis* 18: 432–42.
- de Koning W, van Dam K. 1992. A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Anal Biochem* 204: 118–23.
- Delgado J, Liao JC. 1992. Metabolic control analysis using transient metabolite concentrations. determination of metabolite concentration control coefficients. *Biochem J* 285: 965–72.
- Eriksson P, Andre L, Ansell R, Blomberg A, Adler L. 1995. Cloning and characterization of GPD2, a second gene encoding sn-glycerol 3-phosphate dehydrogenase (NAD⁺) in *Saccharomyces cerevisiae*, and its comparison with GPD1. *Mol Microbiol* 17: 95–107.
- Fiehn O. 2002. Metabolomics - the link between genotypes and phenotypes. *Plant Mol Biol* 48: 155–71.

- Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L. 2000. Metabolite profiling for plant functional genomics. *Nat Biotechnol* 18: 1157–61.
- Fischer E, Sauer U. 2003. Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS. *Eur J Biochem* 270: 880–91.
- Foster J, Gombert AK, Nielsen J. 2002. A functional genomics approach using metabolomics and *in silico* pathway analysis. *Biotechnol Bioeng* 79: 703–12.
- Gombert AK, Moreira dos Santos M, Christensen B, Nielsen J. 2001. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol* 183: 1441–51.
- Gonzalez B, Francois J, Renaud M. 1997. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast* 13: 1347–55.
- Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. 2004. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol* 22: 245–52.
- Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17: 994–9.
- Hatzimanikatis V, Bailey JE. 1996. Effects of spatiotemporal variations on metabolic control: approximate analysis using (log) linear kinetic models. *Biotechnol Bioeng* 54.
- Haunschild MD, Freisleben B, Takors R, Wiechert W. 2005. Investigating the dynamic behavior of biochemical networks using model families. *Bioinformatics* 21: 1617–25.
- Haurie V, Perrot M, Mini T, Jenö P, Sgallio F, Boucherie H. 2001. The transcriptional activator Cat8p provides a major contribution to the reprogramming of carbon metabolism during the diauxic shift in *Saccharomyces cerevisiae*. *J Biol Chem* 276: 76–85.
- Heijnen JJ, Van Gulik WM, Shimizu H, Stephanopoulos G. 2004. Metabolic flux control analysis of branch points: an improved approach to obtain flux control coefficients from large perturbation data. *Metab Eng* 6: 391–400.
- Heinrich R, Rapoport SM. 1977. Metabolic regulation and mathematical models. *Prog. Biophys. Molec. Biol.* 32: 1–82.
- Hofmeyr JH, Cornish-Bowden A. 1996. Co-response analysis: a new experimental strategy for metabolic control analysis. *J Theor Biol* 182: 371–80.
- Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L. 2001. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292: 929–34.
- Kacser H, Burns JA. 1973. The control of flux. *Symp Soc Exp Biol* 27: 65–104.
- King RD, Whelan KE, Jones FM, Reiser PG, Bryant CH, Muggleton SH, Kell DB, Oliver SG. 2004. Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* 427: 247–52.
- Knowles MR, Cervino S, Skynner HA, Hunt SP, de Felipe C, Salim K, Meneses-Lorente G, McAllister G, Guest PC. 2003. Multiplex proteomic analysis by two-dimensional differential in-gel electrophoresis. *Proteomics* 3: 1162–71.
- Kobe B, Kemp BE. 1999. Active site-directed protein regulation. *Nature* 402: 373–6.
- Kresnowati MT, van Winden WA, Heijnen JJ. 2005. Determination of elasticities, con-

- centration and flux control coefficients from transient metabolite data using linlog kinetics. *Metab Eng* 7: 142–53.
- Kretschmer M, Fraenkel DG. 1991. Yeast 6-phosphofructo-2-kinase: sequence and mutant. *Biochemistry* 30: 10663–72.
- Kromer JO, Sorgenfrei O, Klopprogge K, Heinzle E, Wittmann C. 2004. In-depth profiling of lysine-producing *Corynebacterium glutamicum* by combined analysis of the transcriptome, metabolome, and fluxome. *J Bacteriol* 186: 1769–84.
- Lange HC, Eman M, van Zuijlen G, Visser D, van Dam JC, Frank J, de Mattos MJ, Heijnen JJ. 2001. Improved rapid sampling for *in vivo* kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 75: 406–15.
- Mashego MR, van Gulik WM, Vinke JL, Heijnen JJ. 2003. Critical evaluation of sampling techniques for residual glucose determination in carbon-limited chemostat culture of *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 83: 395–9.
- Monod J, Wyman J, Changeux JP. 1965. On the nature of allosteric transitions: A plausible model. *J Mol Biol* 12: 88–118.
- Nielsen J. 1997. Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates. *Biochem J* 321: 133–8.
- Nissler K, Otto A, Schellenberger W, Hofmann E. 1983. Similarity of activation of yeast phosphofructokinase by AMP and fructose-2,6-bisphosphate. *Biochem Biophys Res Commun* 111: 294–300.
- Pahlman AK, Granath K, Ansell R, Hohmann S, Adler L. 2001. The yeast glycerol 3-phosphatases Gpp1p and Gpp2p are required for glycerol biosynthesis and differentially involved in the cellular responses to osmotic, anaerobic, and oxidative stress. *J Biol Chem* 276: 3555–63.
- Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG. 2001. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* 19: 45–50.
- Rizzi M, Baltés M, Theobald U, Reuss M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol Bioeng* 55: 592–608.
- Rohwer JM, Schuster S, Westerhoff HV. 1996. How to recognize monofunctional units in a metabolic system. *J Theor Biol* 179: 213–28.
- Savageau MA. 1976. *Biochemical system analysis*. Addison-wesley, Reading (MA).
- Schaefer U, Boos W, Takors R, Weuster-Botz D. 1999. Automated sampling device for monitoring intracellular metabolite dynamics. *Anal Biochem* 270: 88–96.
- Smedsgaard J, Nielsen J. 2005. Metabolite profiling of fungi and yeast: from phenotype to metabolome by MS and informatics. *J Exp Bot* 56: 273–86.
- Soga T, Ueno Y, Naraoka H, Ohashi Y, Tomita M, Nishioka T. 2002. Simultaneous determination of anionic intermediates for bacillus subtilis metabolic pathways by capillary electrophoresis electrospray ionization mass spectrometry. *Anal Chem* 74: 2233–9.
- Teusink B, Baganz F, Westerhoff HV, Oliver SG. 1998. Metabolic control analysis as a tool in the elucidation of the function of novel genes. *Methods Microbiol.* 26: 297–336.
- Theobald U, Mailinger W, Baltés M, Reuss M, Rizzi M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng* 55: 305–16.

- Theobald U, Mailinger W, Reuss M, Rizzi M. 1993. *In vivo* analysis of glucose-induced fast changes in yeast adenine nucleotide pool applying a rapid sampling technique. *Anal Biochem* 214: 31–7.
- van Dam JC, Eman M, Frank J, Lange HC, van Dedem GW, Heijnen JJ. 2002. Analysis of glycolytic intermediates in *Saccharomyces cerevisiae* using anion exchange chromatography and electrospray ionization with tandem mass spectrometric detection. *Anal Chim Acta* 460: 209–18.
- van Winden WA, van Dam JC, Ras C, Kleijn RJ, Vinke JL, van Gulik WM, Heijnen JJ. 2005. Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of ^{13}C -labeled primary metabolites. *FEMS Yeast Res* 5: 559–68.
- Vaseghi S, Baumeister A, Rizzi M, Reuss M. 1999. *In vivo* dynamics of the pentose phosphate pathway in *Saccharomyces cerevisiae*. *Metab Eng* 1: 128–40.
- Visser D, Heijnen JJ. 2003. Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metab Eng* 5: 164–76.
- Visser D, van Zuylen GA, van Dam JC, Eman MR, Proll A, Ras C, Wu L, van Gulik WM, Heijnen JJ. 2004. Analysis of *in vivo* kinetics of glycolysis in aerobic *Saccharomyces cerevisiae* by application of glucose and ethanol pulses. *Biotechnol Bioeng* 88: 157–67.
- Voit EO. 2000. Computational analysis of biochemical systems: a practical guide for biochemists and molecular biologists. Cambridge University Press, Cambridge.
- Westerhoff HV, van Dam K. 1987. Thermodynamics and control of biological free energy transduction. Elsevier, Amsterdam.
- Wiechert W. 2001. ^{13}C metabolic flux analysis. *Metab Eng* 3: 195–206.
- Wiechert W. 2002. Modeling and simulation: tools for metabolic engineering. *J Biotechnol* 94: 37–63.
- Wiechert W, Mollney M, Petersen S, de Graaf AA. 2001. A universal framework for ^{13}C metabolic flux analysis. *Metab Eng* 3: 265–83.
- Wittmann C, Heinze E. 2001. Application of MALDI-TOF MS to lysine-producing *Corynebacterium glutamicum*: a novel approach for metabolic flux analysis. *Eur J Biochem* 268: 2441–55.
- Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, van Winden WA, van Gulik WM, Heijnen JJ. 2005. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using fully $\text{U-}^{13}\text{C}$ -labeled cell extracts as internal standards. *Anal Biochem* 336: 164–71.
- Wu L, Wang W, van Gulik WM, Heijnen JJ. 2004. A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. *Eur J Biochem* 271: 3348–59.

Appendix

Michaelis-Menten kinetics for reversible reactions with one substrate and one product were used for the lumped enzymes ‘HK’, ‘ENO’, ‘GPD’ and ‘GPP’:

$$v = v_f \frac{\frac{S}{K_s} \left(1 - \frac{P/S}{K_{eq}}\right)}{1 + \frac{S}{K_s} + \frac{P}{K_p}}$$

For fructose 1,6-bisphosphate aldolase (FBA), Michaelis-Menten kinetics for reversible reactions with one substrate and two products was used:

$$v_{\text{FBA}} = v_{f,\text{FBA}} \frac{\frac{S}{K_s} \left(1 - \frac{P_1 P_2 / S}{K_{eq}}\right)}{\left(1 + \frac{S}{K_s}\right) \left(1 + \frac{P_1}{K_{p,1}} + \frac{P_2}{K_{p,2}}\right)}$$

For all these reactions the Michaelis-Menten constants for substrates and products were set to 1 mM. The maximum forward velocities used for these (lumped) enzymes were:

$$\begin{aligned} v_{f,\text{HK}} &= 1 \mu\text{mol/l/min}, \\ v_{f,\text{FBA}} &= v_{f,\text{ENO}} = v_{f,\text{GPD}} = v_{f,\text{GPP}} = 5 \mu\text{mol/l/min} \end{aligned}$$

The equilibrium constants used for these (lumped) enzymes were:

$$K_{eq,\text{HK}} = K_{eq,\text{ENO}} = K_{eq,\text{GPD}} = K_{eq,\text{GPP}} = 1, K_{eq,\text{FBA}} = 0.1 \text{ mM}$$

For the enzyme 6PF-1-K, Monod, Wyman and Changeux kinetics were used (Monod et al., 1965). We have included the stimulatory effect of F2,6bP, assuming that it increases both the affinity of 6PF-1-K for F6P and the maximum velocity of the enzyme, which has been described for the yeast enzyme (Nissler et al., 1983):

$$\begin{aligned} v &= v_{6\text{PF-1-K}}^{\max} [\text{F2,6bP}] \frac{[\text{F6P}]/K_{\text{F6P}}}{1 + [\text{F6P}]/K_{\text{F6P}}} \frac{1}{1 + L_{6\text{PF-1-K}}} \\ L_{6\text{PF-1-K}} &= L_0 \frac{1}{(1 + [\text{F6P}]/K_{\text{F6P}})^{n_{6\text{PF-1-K}}}}, \quad K_{\text{F6P}} = \frac{K_{0.5}}{[\text{F2,6bP}]} \end{aligned}$$

with $K_{0.5} = 1 \text{ mM}$, $n_{6\text{PF-1-K}} = 8$, $L_0 = 10^4$ and $v_{6\text{PF-1-K}}^{\max} = 1 \mu\text{mol/l/min}$. For the lumped step ‘PK’ we used the rate equation for pyruvate kinase, which includes the stimulatory effect of F1,6bP that turns the Hill type kinetics of PK towards PEP into hyperbolic, Michaelis-Menten type kinetics:

$$v = v_{\text{PK}}^{\max} \frac{\frac{[\text{PEP}]}{K_{\text{PEP}}} \left(\frac{[\text{PEP}]}{K_{\text{PEP}}} + 1\right)^{n_{\text{PK}}-1}}{L_{\text{PK}} \left(\frac{2}{1 + [\text{F1,6bP}]/K_{\text{F1,6bP}}}\right)^{n_{\text{PK}}} + \left(\frac{[\text{PEP}]}{K_{\text{PEP}}} + 1\right)^{n_{\text{PK}}}}$$

Chapter 5

with $K_{\text{PEP}} = K_{\text{F1,6bP}} = 1$ mM, $n_{\text{PK}'} = 4$, $L_{\text{PK}'} = 100$ and $v_{\text{PK}'}^{\text{max}} = 10$ $\mu\text{mol/l/min}$. For simplicity reasons, conserved moieties such as ATP, ADP, phosphate, NAD^+ and NADH were not included in the rate equations.

Chapter 6

In Vivo Kinetic Analysis in Prolonged Chemostats

Abstract In this chapter, prolonged chemostat cultivation is applied to investigate *in vivo* enzyme kinetics of *S. cerevisiae*. *S. cerevisiae* was grown in carbon-limited aerobic chemostats for 70–95 generations, during which multiple steady states were observed, characterized by constant intracellular fluxes but significant changes in intracellular metabolite concentrations and enzyme capacities. We provide evidence for two relevant kinetic mechanisms for sustaining constant fluxes: *in vivo* near-equilibrium of reversible reactions and tight regulation of irreversible reactions by coordinated changes of metabolic effectors. Using linear-logarithmic kinetics, we illustrate that these multiple steady state measurements provide linear constraints between elasticity parameters instead of their absolute values. Upon perturbation by a glucose pulse, glucose uptake and ethanol excretion in prolonged cultures were remarkably lower, compared to a reference culture perturbed at 10 generations. Metabolome measurements during the transient indicate that the differences might be due to a reduced ATP regeneration capacity in prolonged cultures.

6.1 Introduction

Directed improvement of cell functionality by genetic interventions relies increasingly on quantitative understanding of complex interactions between different cellular entities, such as the transcriptome, proteome, metabolome and fluxome. In particular, the relationship between fluxes (v), enzyme capacities (e) and metabolite concentrations (\mathbf{x}) can be defined by (nonlinear) enzyme kinetics with a set of parameters (\mathbf{P}):

$$v = f(e, \mathbf{x}, \mathbf{P}) \quad (6.1)$$

Accepted for publication in *Metab Eng* as: Wu L, Mashego MR, Proell AM, Vinke JL, Ras C, van Dam JC, van Winden WA, van Gulik WM, Heijnen JJ. *In vivo* kinetics of primary metabolism in *S. cerevisiae* studied through prolonged chemostat cultivation.

Kinetic parameters can be estimated from multiple (quasi) steady state experiments, in which fluxes, enzyme capacities and metabolite concentrations are determined, as demonstrated for an *in vitro* reconstituted pathway (Giersch, 1995; Wu et al., 2004). Multiple steady states can also be established *in vivo*, by applying e.g. selective inhibition, different physiological conditions and genetic modification. It has been shown that data obtained from such experiments lead to quantitative insights in *in vivo* kinetics (Heijnen et al., 2004; Shimizu et al., 2003; Stephanopoulos and Simpson, 1997; van Gulik et al., 2000). *In vivo* kinetic information can also be obtained from stimulus response experiments, in which a reference (quasi) steady state is perturbed by an external stimulus, e.g. a substrate pulse (Rizzi et al., 1997; Theobald et al., 1997; Visser et al., 2004b; Wu et al., 2003). Within a short transient thereafter (e.g. seconds to a few minutes), the enzyme capacities can be assumed constant due to the relatively slow protein synthesis. Thus, only the measurements of the short-term metabolite responses are required to obtain *in vivo* kinetic parameters.

In both types of approaches, a well-defined (quasi) steady state is crucial to obtain reproducible *in vivo* data. To this end a chemostat has been the preferred tool for cultivation, in which the growth rate as well as the extracellular environment (e.g. residual nutrient concentrations) are subject to tight control (Novick and Szilard, 1950). However, it has been shown that a low residual concentration of the growth-limiting substrate creates a continuous selective pressure, leading to adaptation of cells in prolonged chemostat cultures. For *Saccharomyces cerevisiae*, prolonged glucose-limited chemostat cultivation results in increased affinity for glucose (Adams et al., 1985; Brown et al., 1998) and a decreased residual glucose concentration (Ferea et al., 1999; Mashego et al., 2005). The evolved strains become more elongated, which increases the surface to volume ratio (Adams et al., 1985; Brown and Hough, 1965). In addition, genome rearrangements and altered gene expression have been observed for adaptive clones isolated from prolonged chemostat cultures (Brown et al., 1998; Dunham et al., 2002; Ferea et al., 1999).

Recently, intracellular metabolite concentrations and *in vitro* enzyme activities have been determined during prolonged chemostat cultivation of *S. cerevisiae*. In contrast to constant extracellular fluxes and constant biomass yield on glucose, remarkable decreases of most intracellular metabolite concentrations have been observed, accompanied by a strong decrease of *in vitro* activities of most enzymes in glycolysis that catalyze reversible reactions (Mashego et al., 2005). In view of the vast time-span of the cultivation (i.e. 100 – 200 generations), prolonged chemostats of *S. cerevisiae* can be considered equivalent to a set of multiple (quasi) steady states with varying metabolite concentrations and enzyme capacities, but constant extracellular fluxes. Moreover, due to the difference in enzyme and metabolite makeup of these (quasi) steady states, different dynamic responses can be expected if an external stimulus is applied. As such, prolonged chemostats are able to provide a unique combination of both steady state and dynamic information on *in vivo* kinetics.

In this work, we cultivated *S. cerevisiae* for 70 – 95 generations in aerobic carbon-limited chemostat cultures. The prolonged cultures were subsequently perturbed by glucose pulses, which led to distinct short-term dynamic responses, compared to a non-prolonged reference chemostat culture perturbed at 10 generations. The observations from multiple (quasi) steady states during the prolonged cultivations and from dynamic experiments are used to obtain insight into *in vivo* equilibria and

allosteric mechanisms in the primary metabolism of *S. cerevisiae*.

6.2 Materials and methods

Strain and chemostat cultivation The haploid, wild type *S. cerevisiae* strain CEN.PK 113-7D was grown overnight in shake flask precultures (0.1 l) before being inoculated in a 7 l fermentor (Applikon, Schiedam, The Netherlands). The preculture medium was based on Verduyn et al. (1992) with 15 g/l glucose. The medium used for chemostat cultures was according to Lange et al. (2001) with 27.1 g/l glucose and 1.42 g/l ethanol to obtain a biomass concentration of approximately 15 g_{DW}/l in the steady state. The aerobic, carbon-limited chemostat cultivation was run at a dilution rate of 0.05 h⁻¹. Other conditions applied were: a working volume of 4 l, a pH controlled at 5, a temperature controlled at 30°C, an overpressure of 0.3 bar, a stirrer speed of 600 rpm and a gas feed (air) flow rate of 3 l/min. The dissolved oxygen tension (DOT) was measured by a Mettler-Toledo DOT sensor (Mettler-Toledo GmbH, Switzerland) and was around 70% in the steady state. The fermentor off-gas was cooled in a condenser connected to a cryostat at 2°C and dried with a Perma Pure dryer (Perma Pure Inc, Toms River, NJ, USA). The O₂ and CO₂ content in the off-gas were subsequently measured by a NGA200 gas analyzer (Rosemount Analytics, Santa Clare, CA, USA).

Determination of culture dry weight Culture samples (5 ml, in triplicate) were filtered through predried and preweighed nitrocellulose filters (pore size 0.45 μm, Gelman Science, Ann Arbor, MI, USA). The filters were washed twice with 5 ml demineralized water, dried at 70°C for 48 hours and weighed.

Extracellular metabolite concentrations Culture supernatant was sampled using pre-cooled stainless steel beads (Mashego et al., 2003). The concentration of glucose and acetate in the supernatant was analyzed with Enzytec™ kits (kit no. 1002781 for glucose and 1002811 for acetate, Scil Diagnostics GmbH, Martinsried, Germany). The concentration of ethanol in the supernatant was analyzed with Enzymatic Bioanalysis kit (kit no. 10176290035, R-Biopharm AG, Darmstadt, Germany). Absorbance was read on an Agilent 8453-UV-visible spectroscopy system (Agilent Technologies Deutschland GmbH, Waldbronn, Germany).

Intracellular metabolite concentrations Rapid sampling and metabolite extraction of biomass (in approximately 1 ml of culture broth) was performed as described previously (Mashego et al., 2005). The concentrations of G6P, F6P, F1,6bP, 2PG, 3PG, PEP in the cell extract was analyzed by LC-ESI-MS/MS (van Dam et al., 2002). F2,6bP, G1P, T6P, 6PG, Pyr, Cit, iCit, αKG, Fum, Mal and Pi were analyzed with the same analysis method and conditions. The metabolites 2PG and 3PG could not be resolved with the applied analytical procedure, the same held for Cit and iCit. Consequently, only the sum of these compounds was determined. For F2,6bP, relative peak areas were measured instead of absolute concentrations.

ATP concentration in the cell extract was analyzed by the ATP Bioluminescence assay kit CLS II (kit no. 1699695, Roche Diagnostics GmbH, Mannheim, Germany). Luminescence was read on a Mediators PhL plate reader (Mediators Diagnostics

GmbH, Vienna, Austria).

Glucose pulse experiment Under steady state conditions, 10 ml of a glucose solution (200 g/l in a syringe was injected pneumatically into the fermentor to increase the residual glucose concentration by 0.5 g/l. Samples for the extra- and intracellular metabolite analysis were taken both in the steady state (i.e. prior to the glucose pulse) and during the transient.

Metabolic flux analysis Intracellular fluxes were estimated by conventional metabolic flux analysis (Stephanopoulos et al., 1998) using steady state measurements reconciled according to van der Heijden et al. (1994). A biomass composition determined for a growth rate of 0.05 h^{-1} (Lange and Heijnen, 2001) and a noncompartmented network of *S. cerevisiae* was used (Stuckrath et al., 2002). Acetaldehyde dehydrogenase (AADH) was assumed NADP^+ -specific and isocitrate dehydrogenase NAD^+ -specific. The inclusion of AcCoA synthetase creates a parallel route for the synthesis of AcCoA from pyruvate, which can proceed either via pyruvate dehydrogenase or via the pyruvate bypass, by the consequent actions of pyruvate decarboxylase, AADH and AcCoA synthetase. In the applied network, NADPH needed for growth is regenerated by the pentose phosphate (PP) pathway and the NADP^+ -AADH. Hence, intracellular fluxes can be made observable by specifying the PP pathway flux (or equivalently, the PP split ratio). Since NADP^+ -AADH is also responsible for metabolizing the ethanol present in the glucose/ethanol mixed substrate feed used in this study, we assumed that the specific NADPH regeneration ($\text{mmol}_{\text{NADPH}}/\text{g}_{\text{DW}}$) via the PP pathway decreases linearly with increasing ethanol fraction in a glucose/ethanol mixed substrate feed, up to an ethanol fraction of 0.4 $\text{Cmol}_{\text{EtOH}}/\text{Cmol}_{\text{Carbon}}$. Above this limit, the metabolic network would change due to the induction of gluconeogenic enzymes (de Jong-Gubbels et al., 1995; Vanrolleghem et al., 1996). For the *Saccharomyces* strain used in this study, the PP split ratio has been estimated to be 0.44 and 0.24 from two ^{13}C -labeling studies (Gombert et al., 2001; van Winden et al., 2005), for ethanol fractions of 0 and 0.28 Cmol/Cmol respectively at a growth rate of 0.1 h^{-1} . For the feed used in this study with an ethanol fraction of 0.17 Cmol/Cmol , a PP split ratio of 0.31 is found, with which the intracellular fluxes were subsequently estimated.

6.3 Results and discussion

Characterization of multiple steady states during prolonged cultivation

A reference chemostat (R) and two prolonged chemostats (PA and PB) of *S. cerevisiae* were carried out at a dilution rate of 0.05 h^{-1} under aerobic carbon-limited conditions. Chemostat PB is identical to chemostat 2 described in (Mashego et al., 2005).

Table 6.1 summarizes (quasi) steady state extracellular measurements at different culture ages. No significant changes are observed in the specific conversion rates of glucose, O_2 , CO_2 and biomass, while the residual concentration of glucose drops. Intracellular metabolite concentrations decrease as well, as shown in Table 6.2 for chemostat PB at 10 and 95 generations, respectively. In addition, Mashego et al. (2005) reported *in vitro* enzyme activity changes in prolonged chemostats. These

Table 6.1 Steady state physiology of chemostat cultures

| Chemostat | Quasi steady state | Generations | c_x | q_{O_2} | q_{CO_2} | q_s | c_s |
|-----------|--------------------|-------------|-------|-------------------------|------------|-------|-------|
| | | | g/l | mmol/g _{DW} /h | | | mM |
| R | R10 | 10 | 14.2 | 1.7 | 1.6 | 0.59 | 0.15 |
| PA | PA10 | 10 | 14.9 | 1.6 | 1.6 | 0.56 | ND |
| | PA70 | 70 | 14.6 | 1.6 | 1.5 | 0.57 | 0.07 |
| PB | PB10 | 10 | 14.9 | 1.5 | 1.4 | 0.56 | 0.11 |
| | PB95 | 95 | 14.6 | 1.5 | 1.5 | 0.57 | 0.04 |

c_x , biomass concentration; q_{O_2} , specific O₂ uptake rate; q_{CO_2} , specific CO₂ evolution rate; q_s , specific substrate uptake rate; c_s , residual glucose concentration.

are given in Table 6.2 for completeness.

The intracellular fluxes were estimated with conventional metabolic flux analysis, assuming that the metabolic network and the biomass composition remain unchanged during the prolonged cultivation.

Mashego et al. (2005) observed that *in vitro* activities of isocitrate lyase and malate synthase gradually increased with culture age. The two enzymes participate in the glyoxylate cycle, which has not been included in the metabolic network used for flux estimation. The glyoxylate cycle is required for growth at an ethanol/glucose

Table 6.2 Intracellular metabolite concentrations ($\mu\text{mol/g}_{\text{DW}}$) and *in vitro* enzyme activities (U/mg prot) in chemostat PB

| Metabolite | Quasi steady state | | Enzyme | Quasi steady state | |
|-------------|--------------------|------|--------|--------------------|------------------|
| | PB10 | PB95 | | PB10 | PB95 |
| G6P | 2.1 | 0.35 | HK | 2.8 | 3.1 |
| F6P | 0.45 | 0.09 | PGI | 4.6 | 2.8 |
| F1,6bP | 0.19 | 0.11 | PFK | 0.3 | 0.3 |
| 2PG+3PG | 1.2 | 0.07 | FBA | 1.5 | 0.5 |
| PEP | 1.1 | 0.17 | TPI | 126 | 29 ^a |
| Pyr | 0.06 | 0.03 | GAPDH | 6.5 | 1.4 |
| Cit+iCit | 4.4 | 0.71 | PGK | 14 | 1.0 |
| α KG | 0.07 | 0.02 | PGM | 11 | 3.7 |
| Suc | 0.06 | 0.03 | ENO | 1.0 | 0.2 ^a |
| Fum | 0.06 | 0.02 | PK | 3.5 | 2.8 |
| Mal | 0.36 | 0.14 | G6PDH | 0.5 | 1.8 |
| G1P | 0.34 | 0.13 | PDC | 0.6 | 0.4 ^a |
| T6P | 0.34 | 0.21 | ADH | 13 | 13 ^a |
| 6PG | 0.39 | 0.23 | | | |
| ATP | 7.8 | 5.8 | | | |
| ADP | 3.8 | 2.3 | | | |
| Pi | 44 | 44 | | | |

^a Determined at 74 generations instead of 95 generations.

ratio above 0.4 Cmol/Cmol in the feed (de Jong-Gubbels et al., 1995), which is considerably higher than the ethanol fraction used in this study (i.e. 0.07 Cmol/Cmol). Moreover, *in vitro* activities of the two enzymes remained very low during the prolonged cultivation; even at PB95 these are much lower than the *in vitro* activities corresponding to a glyoxylate cycle requirement (de Jong-Gubbels et al., 1995). We therefore conclude that the metabolic network applied here (i.e. without the glyoxylate cycle) is adequate for flux estimation in the prolonged cultivation.

The constant extracellular fluxes also point to an unchanged biomass composition, although large decreases in *in vitro* activities of glycolytic enzymes might imply decreased abundance of these enzymes. During the prolonged cultivation, a constant rate of base addition was observed (not shown). This corresponds to a constant ammonium uptake rate, as base was added to the fermentor to compensate pH decrease due to the uptake of ammonium, the sole nitrogen source. As proteins contain up to 92% of the cellular nitrogen (calculated according to the biomass composition in Lange and Heijnen, 2001), a constant ammonium uptake indicates an unchanged biomass protein content.

Intracellular fluxes obtained from metabolic flux analysis are depicted in Fig. 6.1 for the reference chemostat at 10 generations. Nearly identical flux distributions were found for different culture ages for the prolonged chemostats (not shown), due to practically unchanged extracellular fluxes. The multiple steady states occurring during the prolonged cultivation can thus be characterized by a set of constant intracellular fluxes and significantly changed enzyme activities and metabolite levels.

Kinetic mechanisms responsible for constant fluxes in multiple steady states The Flux through an enzyme is determined by the *in vivo* enzyme capacity and its modification by metabolic effectors. Multiple steady state measurements of fluxes, *in vitro* enzyme activities and metabolite concentrations from prolonged chemostat cultures essentially provide information of *in vivo* enzyme/metabolite interactions. We have assumed here that changes in *in vitro* enzyme activities correspond to changes in *in vivo* enzyme capacities.

A possible explanation of the apparent constant fluxes, in contrast to large changes in enzyme and metabolite levels among multiple steady states, would be an increased affinity of enzymes for their substrates. However, this would require point mutations occur to a large number of enzymes, due to the widespread decrease in metabolite concentrations, which seems highly unlikely. Instead, enzyme activity changes follow a clear pattern during the prolonged cultivation. *In vitro* activities of enzymes catalyzing irreversible steps in glycolysis, i.e. HK, PFK and PK (Voet and Voet, 2004), remained rather constant, while activities of enzymes catalyzing reversible reactions, i.e. PGI, FBA, TPI, GAPDH, PGK, PGM and ENO, all decreased during the prolonged cultivation. The systematic changes indicate gene regulatory events on enzyme levels rather than alterations of individual enzyme kinetic properties.

Mashego et al. (2005) suggested that the reversible enzymes are close to equilibrium *in vivo*, since *in vivo* mass action ratios of metabolites involved in (lumped) reversible reactions are rather constant during the prolonged cultivation. Hence, decreasing enzyme capacities will not affect the fluxes through these enzymes. This is supported by the very similar labeling distributions of intracellular metabolites

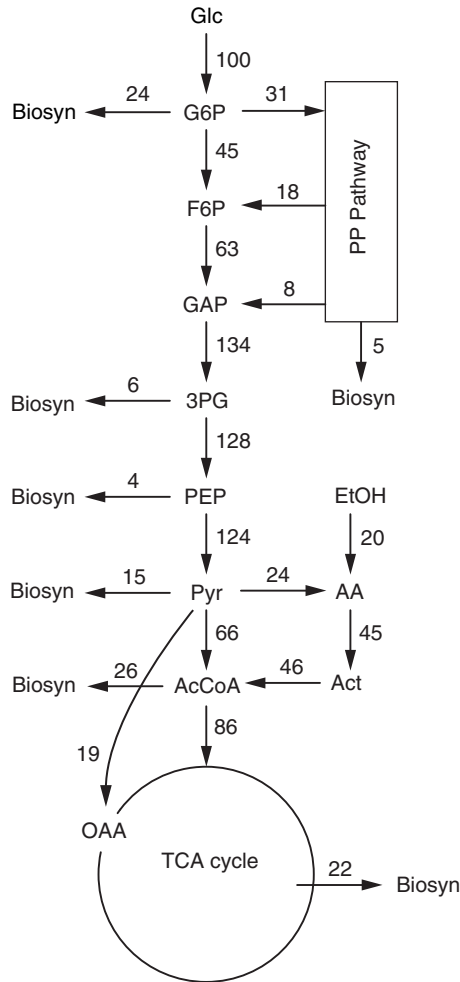


Figure 6.1 Estimated intracellular fluxes (normalized upon the glucose influx) in the reference chemostat R at 10 generations.

involved in highly reversible reactions, measured during ^{13}C -labeling experiments in a 10-generation-old chemostat culture of *S. cerevisiae* (van Winden, 2002; van Winden et al., 2005). Moreover, enzymes catalyzing reversible reactions are apparently present in overcapacity during the prolonged cultivation. Table 6.3 shows that with exception of ENO, all measured *in vitro* maximum activities largely surpass the estimated fluxes through these enzymes.

The rather constant enzyme capacities of HK, PFK and PK imply that different kinetic mechanisms are responsible for irreversible reactions, which are discussed in detail below.

Irreversible enzymes are regulated by coordinated changes of metabolic effectors and enzyme levels The *in vitro* enzyme activities of HK, PFK and PK, which catalyze irreversible steps in glycolysis, remained more or less constant

Table 6.3 Ratio of *in vitro* activities to the estimated fluxes through reversible steps in glycolysis

| Enzyme | PB10 | PB95 |
|--------|------|------------------|
| PGI | 113 | 68 |
| FBA | 37 | 12 |
| TPI | 3096 | 705 ^a |
| GAPDH | 160 | 35 |
| PGK | 332 | 25 |
| PGM | 265 | 90 |
| ENO | 25 | 6 ^a |

Calculated assuming the release of 53% of total protein in cell extracts used for enzyme activity assays (Jansen ML, personal communication) and a cellular protein content according to Lange and Heijnen (2001).

^a Determined at 74 generations instead of 95 generations.

through the prolonged cultivation. In addition, concentrations of the adenine nucleotides that are involved in all these reactions, as well as the energy charge, do not significantly change (Mashego et al., 2005). Constant intracellular fluxes agree with these observations, but not with drastic fold changes of other metabolites involved in these irreversible reactions. Table 6.4 summarizes major effectors of these enzymes other than adenine nucleotides, along with their concentration changes between quasi steady state PB10 and PB95. It can be seen that decrease of concentrations of positive effectors (i.e. substrate and allosteric activators) is accompanied by decrease of concentrations of negative effectors (i.e. products and allosteric inhibitors). This suggests that coordinated changes of both positive and negative effectors have as a net effect that intracellular fluxes through these enzymes remain constant.

For HK, the concentration of the substrate intracellular glucose was not measured. However, it is expected to decrease due to the measured decrease in extracellular glucose concentration (Glc^{Ex}), as hexose transport in *S. cerevisiae* is a facilitated diffusion process (Maier et al., 2002). T6P has been shown to strongly inhibit HK *in vitro* (Blazquez et al., 1993). Indeed, Fig. 6.2 shows that the concentration of extracellular glucose, G6P and T6P decreased concomitantly during prolonged

Table 6.4 Concentration changes of metabolic effectors involved in the irreversible reactions in glycolysis

| Enzyme | HK | | | PFK | | | | PK | | |
|-------------|--------------------------|------|------|------|--------|--------|------|------|--------|------|
| | Glc^{Ex} | G6P | T6P | F6P | F1,6bP | F2,6bP | Cit | PEP | F1,6bP | Cit |
| Fold change | -2.4 | -6.1 | -1.6 | -5.0 | -1.6 | -2.4 | -6.2 | -6.5 | -1.6 | -6.2 |
| Effect | + | - | - | + | - | + | - | + | + | - |

Fold changes were calculated based on concentrations measured at 10 generations and 95 generations in chemostat PB. The fold change of Cit was calculated using the measured sum of Cit and iCit concentrations. +: stimulatory effect; -, inhibitory effect.

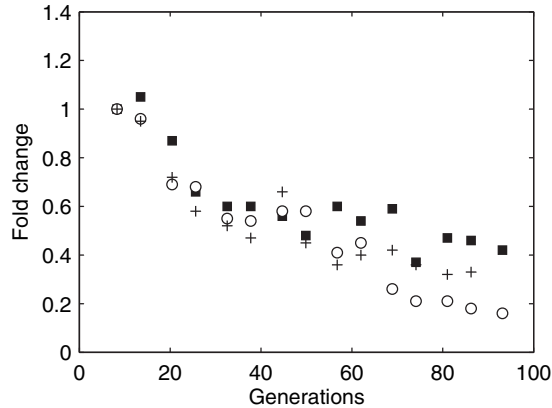


Figure 6.2 Coordinated concentration changes of hexokinase effectors during prolonged chemostat PB. Extracellular glucose: ■; G6P: ○; T6P: +.

chemostat PB.

Yeast PFK is subject to tight allosteric control (Sols, 1981). The principle allosteric mechanism of yeast PFK seems to be the positive cooperative binding of F6P and ATP inhibition. The latter can be counteracted by AMP and F2,6bP (Avigad, 1981; Nissler et al., 1983), two major activators of PFK. F2,6bP activation is however significantly impaired in the presence of F1,6bP *in vitro* (Przybylski et al., 1985; Teusink et al., 2000). Furthermore, feedback inhibition by citrate has been observed for eukaryotic PFK (Li et al., 1999; Parmeggiani and Bowman, 1963; Salas et al., 1965; Yoshino and Murakami, 1982), including that of yeast. During prolonged cultivation, the concentration of adenine nucleotides and inorganic phosphate remained rather constant, while concentrations of both positive effectors (F6P and F2,6bP) and negative effectors (F1,6bP and citrate) decreased, consistent with the known allosteric mechanisms. Moreover, at physiological concentrations of other effectors, yeast PFK has been found sensitive to F6P only above 0.5 mM *in vitro*, due to the cooperative effect of F6P binding (Teusink, 1999). In chemostat PB, the F6P concentration was much lower (0.19 – 0.04 mM, assuming a cytosolic volume of 2.38 ml/g_{DW} (Ditzelmuller et al., 1983)), which implies a limited sensitivity of PFK toward the decreasing F6P concentration.

For yeast pyruvate kinase isoenzyme encoded by *PYK1* (Boles et al., 1997), the major regulatory feature is the Hill type kinetics with respect to PEP and the transition into hyperbolic, Michaelis-Menten type kinetics in the presence of the potent activator F1,6bP (Hess et al., 1966; Murcott et al., 1992). *In vitro* studies show further that PK is activated by ADP and Mg²⁺, inhibited by ATP and citrate (Haecel et al., 1968) and apparently insensitive to pyruvate (Macfarlane and Ainsworth, 1972; Teusink et al., 2000). Table 6.4 shows that concentrations of both positive effectors (PEP, F1,6bP) and negative effectors (citrate) decreased during the prolonged cultivation, consistent with the *in vitro* observed regulatory mechanisms. Alternatively, it cannot be excluded that an increased expression of the F1,6bP-insensitive isoenzyme encoded by *PYK2*, as well as an increased phosphorylation of the PK isoenzymes might also play a role in homeostasis of the flux through PK

(Portela et al., 2002).

For irreversible reactions, there might be other adaptation mechanisms than coordinated changes of effector concentrations. An example is the first step into the PP pathway, catalyzed by G6PDH. It is essentially irreversible due to fast hydrolysis of the reaction product, 6-phospho-D-glucono-1,5-lactone into 6-phosphogluconate (Voet and Voet, 2004). For the same reason, intracellular level of 6-phospho-D-glucono-1,5-lactone might be very low and might not influence the reaction rate. Assuming further that cells would maintain a constant NADPH/NADP⁺ ratio, analogous to the constant energy charge observed in Mashego et al. (2005), G6PDH activity is then only known to be affected by G6P concentration. This explains the observed increase in the *in vitro* G6PDH activity (see Table 6.2), which compensates for the decreasing G6P concentration during the prolonged cultivation.

Derivation of elasticities and elasticity constraints from multiple steady state measurements The multiple steady state measurements obtained from the prolonged cultivation can also provide quantitative *in vivo* kinetic information, next to qualitative insights into *in vivo* equilibria and allosteric interactions. This is illustrated with a simple approximative kinetic formulation, i.e. linear-logarithmic (lin-log) kinetics (Visser and Heijnen, 2002; Visser et al., 2004a; Wu et al., 2004) as an example. In the lin-log kinetic format, the reaction rate v_i through an enzyme i with capacity e_i is given as

$$\frac{v_i}{v_i^0} = \frac{e_i}{e_i^0} \left(1 + \varepsilon_{i,1}^0 \ln \frac{x_1}{x_1^0} + \varepsilon_{i,2}^0 \ln \frac{x_2}{x_2^0} + \dots \right) \quad (6.2)$$

where 0 denotes a reference state and ε stands for elasticity. The elasticity serves as a kinetic parameter, defined as the scaled sensitivity of the reaction rate toward the metabolite concentrations:

$$\varepsilon_{i,j}^0 = \frac{x_j^0}{v_i^0} \left. \frac{\partial v_i}{\partial x_j} \right|_0 \quad (6.3)$$

For near-equilibrium reactions involving two or more reactants, linear constraints between the elasticities have been derived (Visser et al., 2004b). For irreversible reactions in glycolysis, both the enzyme capacities and the intracellular fluxes did not significantly change during the prolonged cultivation. Choosing quasi steady state PB10 as a reference state and taking into account that $v_i/v_i^0 = 1$ and $e_i/e_i^0 = 1$, Eq. 6.2 can be simplified to:

$$\varepsilon_{i,1}^0 \ln \frac{x_1}{x_1^0} + \varepsilon_{i,2}^0 \ln \frac{x_2}{x_2^0} + \dots = 0 \quad (6.4)$$

With concentrations of n effectors measured in m steady states, a set of linear equations like Eq. 6.4 can be formulated in the matrix form:

$$\mathbf{A} \cdot \mathbf{E}^{\mathbf{x}^0} = \mathbf{0} \quad (6.5)$$

where $\mathbf{E}^{\mathbf{x}^0}$ is a $n \times 1$ vector containing the elasticities of each reaction and \mathbf{A} is a $m \times n$ matrix containing the logarithms of fold changes of effector concentrations (i.e. $[\ln \mathbf{x}/\mathbf{x}^0]$). In principle, the elasticities can be solved or estimated from Eq. 6.5 provided that $m \geq n$, as demonstrated previously (Wu et al., 2004). It is however

observed that the metabolome measurements from prolonged chemostats have low information content, which leads to unidentifiability of the elasticities values. The causes are two fold. First, concentrations of some metabolites will not significantly change during the prolonged cultivation, such as the adenine nucleotides. The elasticities of these metabolites can therefore not be obtained. Second, highly coordinated changes of metabolite concentrations (see e.g. Fig. 6.2) can lead to linear dependencies between concentrations measured in different steady states (i.e. $m < n$), which impedes estimation of the absolute values of the elasticities. When constrained linear optimization was applied to estimate the elasticities, it appeared that the obtained elasticity values were highly dependent on initial guesses, while linear relationship exists between elasticity ratios (e.g. $\varepsilon_2/\varepsilon_1$ and $\varepsilon_3/\varepsilon_1$ for an enzyme with 3 effectors) irrespective of the initial guesses used (not shown), indicating significant co-linearity between the steady state measurements. However, the metabolome data can still yield linear algebraic relationship(s) between the individual elasticities, as shown in the Appendix for the irreversible glycolytic reactions. Following the algebraic constraints given in Table 6.5, the one remaining elasticity can be easily calculated if values of two other elasticities are known. Another constraint on the elasticities follows from the definition of elasticities in Eq. 6.3: positive effectors will have positive elasticities and negative effectors will have negative elasticities. For HK for example, it holds that $\varepsilon_{\text{HK,Glc}}^0 > 0$, $\varepsilon_{\text{HK,G6P}}^0 < 0$ and $\varepsilon_{\text{HK,T6P}}^0 < 0$.

Simultaneous changes of enzyme capacities and effector concentrations render improved identifiability of elasticities in prolonged chemostat experiments. For example, the elasticity of G6P for G6PDH is calculated to be 0.42 ($R^2 = 0.96$) directly from Eq. 6.2, assuming that G6P is the only effector with a significantly changed concentration.

The obtained constraints for reversible and irreversible reactions can be combined with measurements from stimulus response experiments. The inclusion of these constraints in the parameter identification procedure, such as implemented in (Kresnowati et al., 2005), leads essentially to a reduction of the number of elasticity parameters that need to be estimated.

Prolonged chemostat cultivation has a pronounced effect on dynamic responses to a glucose pulse The reference chemostat R and the prolonged chemostats PA and PB were perturbed by glucose pulses at 10, 70 and 95 generations, respectively. The dynamic responses are summarized in Fig. 6.3-6.7.

The short-term responses of extracellular metabolite in prolonged cultures differ

Table 6.5 Linear algebraic constraints for elasticities of the irreversible reactions in glycolysis

| | |
|-----|---|
| HK | $\varepsilon_{\text{HK, Glc}}^0 + 1.71 \varepsilon_{\text{HK, G6P}}^0 + 1.40 \varepsilon_{\text{HK, T6P}}^0 = 0$ |
| PFK | $\varepsilon_{\text{PFK, F6P}}^0 + 0.50 \varepsilon_{\text{PFK, F2,6bP}}^0 + 1.46 \varepsilon_{\text{PFK, Cit}}^0 = 0$ $\varepsilon_{\text{PFK, F1,6bP}}^0 + 0.14 \varepsilon_{\text{PFK, F2,6bP}}^0 + 0.87 \varepsilon_{\text{PFK, Cit}}^0 = 0$ |
| PK | $\varepsilon_{\text{PK, PEP}}^0 + 0.38 \varepsilon_{\text{PK, F1,6bP}}^0 + 0.81 \varepsilon_{\text{PK, Cit}}^0 = 0$ |

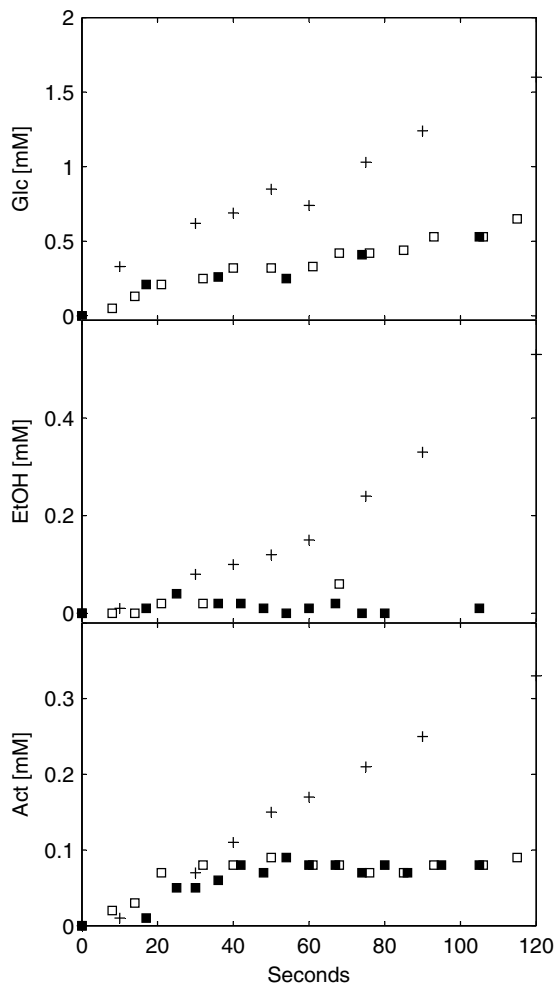


Figure 6.3 Cumulative glucose uptake, ethanol production and acetate production after a glucose pulse given at $t = 0$. Quasi steady states: R10: +, PA70: ■, PB95: □.

significantly from those in the reference culture (6.3). The prolonged chemostats showed a diminished glucose uptake (~ 1.0 mmol/g_{DW}/h) compared to the reference (~ 3.0 mmol/g_{DW}/h) and did not produce any measurable ethanol. Acetate production in the prolonged cultures was comparable to that of the wild type up to 40 seconds after the glucose pulse, but ceased thereafter.

The rapid dynamics of *in vivo* OUR and CER after a glucose pulse can be reconstructed from DOT and off-gas measurement (Bloemen et al., 2003). This is however hampered by the DOT sensor drift after months of chemostat cultivation. We hence compared the relative changes of measured off-gas O₂ and CO₂ concentration (in mol%) compared to the steady state (Fig 6.4). Consistent with a diminished glucose uptake and ethanol/acetate excretion, the O₂ consumption and CO₂ evolution is

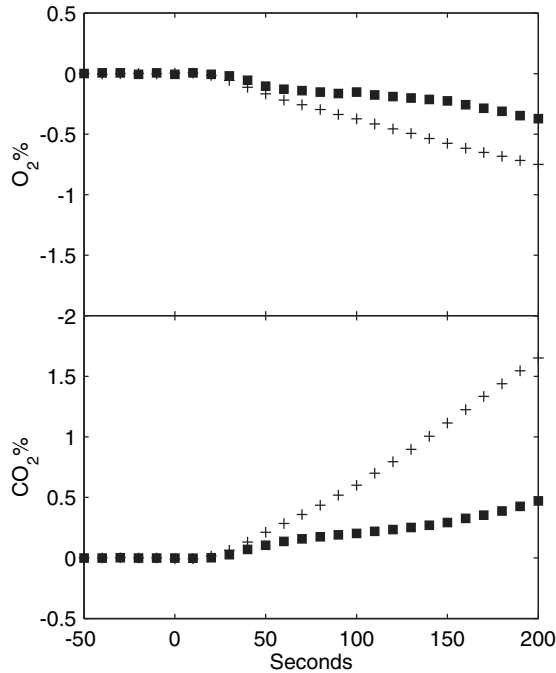


Figure 6.4 Changes of off-gas O_2 and CO_2 concentration (in mol%) relative to the steady state, after a glucose pulse given at $t = 0$. Quasi steady states: R10: +, PA70: ■. Time delay caused by the fermentor off-gas system was corrected according to Bloemen et al. (2003).

much less in chemostat PA compared to the reference.

Fig. 6.5 shows the responses of several glycolytic and TCA cycle intermediates after the glucose pulse. Responses of glycolytic intermediates are much less pronounced in the prolonged cultures compared to the reference, although trends of individual metabolites are similar. The concentration of G6P, F6P (not shown), F1,6bP and T6P all increased due to the increased glucose influx. The concentration of 2PG, 3PG (not shown) and PEP decreased, which has been interpreted as a result of activation of PK by F1,6bP, which drains the triose phosphate pool (Theobald et al., 1997). The concentration of TCA cycle intermediates remained practically unperturbed in the prolonged chemostats, while in the reference chemostat a steady increase of α -KG and Mal was observed, indicating an increased flux into the TCA cycle.

Fig. 6.6 shows the displacement from equilibrium of the reactions catalyzed by PGI, PGM and ENO following the glucose pulse. After an initial decrease, the mass action ratio of PGI quickly relaxes back to the steady state level. This indicates that PGI is also close to equilibrium during the transient, according to the criterion outlined in Reich and Sel'kov (1981). For PGM and ENO however, the combined mass action ratio sharply decreased after the perturbation and did not relax to the steady state value. Thus, in order to accommodate an increased flux through lower glycolysis, these two reactions need to be displaced further away from the equilibrium than in the steady state to create a larger thermodynamic driving force.

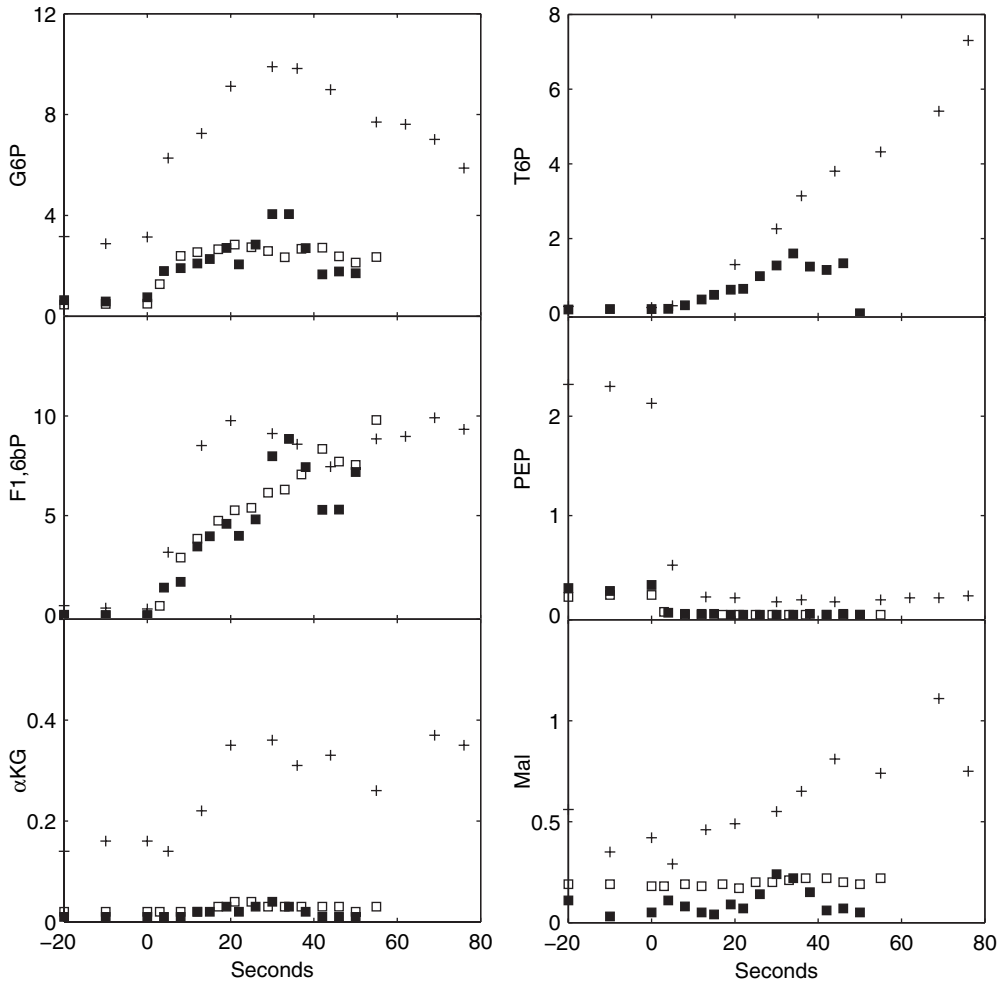


Figure 6.5 Dynamic responses of several representative glycolytic and TCA cycle intermediates after a glucose pulse given at $t = 0$. Quasi steady states: R10: +, PA70: ■, PB95: □. T6P was not measured in chemostat PB.

Fig. 6.7 shows a comparable steady state ATP concentration in the reference and the prolonged chemostat PA. After the glucose pulse, the ATP concentration decreased initially in both cultures, due to increased ATP consumption for phosphorylation. The drop in ATP is however much more pronounced in the prolonged culture, where the ATP concentration remained low after the initial decrease, compared to a rapid recovery seen in the reference chemostat.

The most intriguing short-term response of the prolonged cultures to glucose excess is the decreased glucose uptake and ethanol excretion, compared to the reference culture. This can be elucidated by measured differences of metabolite responses and enzyme capacities between the reference and prolonged cultures.

The continuous generation of ATP is necessary to sustain an increased glycolytic flux after the glucose pulse. However, glycolysis follows a ‘turbo design’ (Teusink

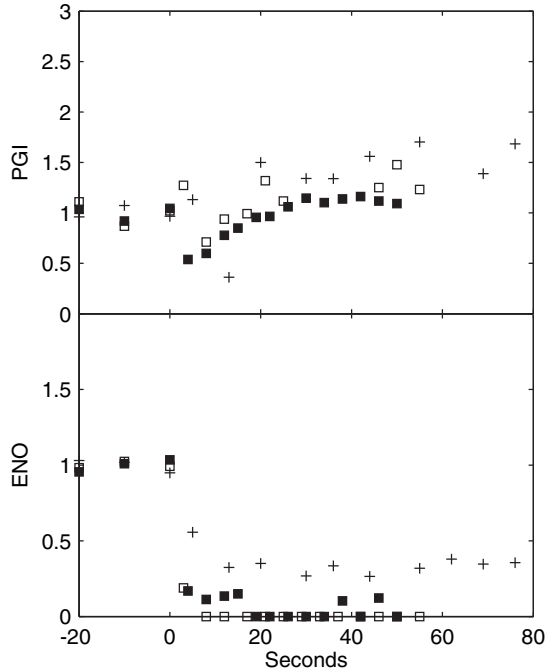


Figure 6.6 Relative change of mass action ratio (with respect to the steady state) after a glucose pulse given at $t = 0$. The mass action ratios are defined according to Quasi steady states: R10: +, PA70: ■, PB95: □.

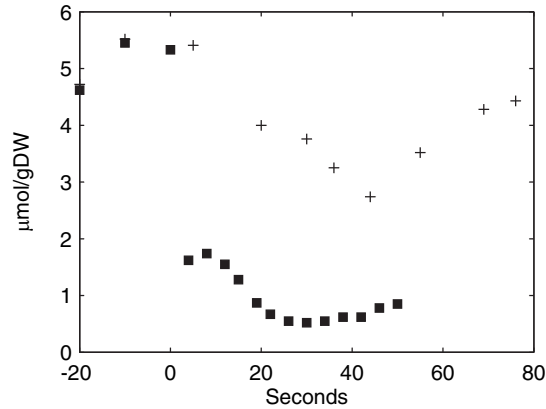


Figure 6.7 Dynamic response of ATP after a glucose pulse given at $t = 0$. Quasi steady states: R10: +, PA70: ■.

et al., 1998): ATP must first be invested in the phosphorylation steps before a surplus of ATP can be yielded. During the transient induced by a glucose pulse, both fermentative and oxidative metabolism generates ATP via substrate phosphorylation in lower glycolysis. With oxidative metabolism, additional ATP is generated by succinyl-CoA synthase and oxidative phosphorylation, where NADH produced in

lower glycolysis and the TCA cycle is oxidized. Hence, a continuous regeneration of ATP is in turn dependent on a continuous supply of glucose into the fermentative and oxidative pathways.

The sharp drop in ATP concentration and the absence of its recovery (Fig. 6.7) observed in the prolonged chemostat PA indicates a shortage of ATP regeneration. This is consistent with decreased fluxes through fermentative and oxidative ATP generation pathways, as can be seen from the near absence of ethanol/acetate production (Fig. 6.3) and a lower O₂ consumption (Fig. 6.4) in the prolonged chemostats. The decreased ATP regeneration in prolonged cultures is likely due to decreased capacities of enzymes in lower glycolysis, which are involved in both energy-providing mechanisms. This is evidenced by a near 100% displacement from equilibrium of the reactions catalyzed by PGM and ENO in the prolonged chemostats following a glucose pulse (Fig. 6.6). In the reference culture however, the two reactions are about 70% displaced from equilibrium, whereas the glucose influx is 3 times larger than the prolonged chemostats.

Mashego et al. (2005) proposed that maintaining high levels of intracellular storage materials (such as trehalose) and overcapacity of enzymes might present a metabolic burden for cells and hence, decreases in both would provide an evolutionary advantage. This seems to hold as well for the overcapacity for ATP regeneration, which had disappeared at the end of the prolonged cultivation, evidenced by the responses to a glucose pulse. Adequate capacity of ATP regeneration and levels of intracellular storage are necessary for cells to respond quickly to sudden changes in the environment. The near absence of such changes in a chemostat thus leads to a population of cells with reduced flexibility to respond to environmental stimuli.

On the other hand, intracellular metabolic intermediates (e.g. glycolytic and TCA cycle intermediates) are already present in low concentrations at 10 generations. Further decreases of the concentrations of these intermediates during prolonged cultivation maintain constant intracellular fluxes through allosteric mechanisms, as discussed before. In addition, decreasing intermediate pools and constant intracellular fluxes lead to a faster turnover of metabolite pools (i.e. a decreased turnover time given by x_j/v_i). This agrees with the general pathway optimization principle in analogy with a factory (Cascante et al., 1996), where an efficient manufacturing process is achieved by minimizing the inventory throughput time (i.e. the inventory divided by sales). Summarizing, the overall result of prolonged chemostat cultivation is an increase in metabolic productivity (characterized by decreased turnover times), at a price of decreased flexibility (characterized by diminishing overcapacities).

6.4 Conclusion

Prolonged chemostat cultivation of *S. cerevisiae* under carbon-limitation leads to the occurrence of multiple steady states, which are characterized by unchanged intracellular fluxes but different metabolite and enzyme levels. In these multiple steady states, two classes of enzymes can be distinguished, where different kinetic mechanisms are involved to sustain a constant intracellular flux. Enzymes catalyzing highly reversible reactions operate close to equilibrium, while enzymes catalyzing irreversible reactions are tightly regulated by coordinated changes of effector con-

centrations and/or enzyme levels. With the lin-log kinetic format, it is shown that measurements obtained from multiple steady states during prolonged cultivation are not sufficient for the identification of individual elasticity parameters. Instead, linear constraints between the elasticities can be obtained. This justifies the application of dynamic perturbation experiments, which yield more *in vivo* kinetic information.

We report, for the first time, dynamic responses of prolonged chemostat cultures of more than 70 generations to a glucose pulse. Compared to a reference chemostat at 10 generations, the prolonged cultures showed decreased glucose uptake and ethanol production. This underlines the importance of culture age in dynamic, stimulus-response experiments conducted in chemostats to investigate *in vivo* kinetics. The difference in dynamic responses between the prolonged and reference culture is likely due to a reduced ATP regeneration capacity after prolonged cultivation, caused by decreased enzyme activities in lower glycolysis. Decrease in overcapacities in cells during prolonged chemostat cultivation is accompanied by faster turnover of metabolite pools, with the overall result of optimized metabolic productivity at the cost of metabolic flexibility.

References

- Adams J, Paquin C, Oeller PW, Lee LW. 1985. Physiological characterization of adaptive clones in evolving populations of the yeast, *Saccharomyces cerevisiae*. *Genetics* 110: 173–85.
- Avigad G. 1981. Stimulation of yeast phosphofructokinase activity by fructose 2,6-bisphosphate. *Biochem Biophys Res Commun* 102: 985–91.
- Blazquez MA, Lagunas R, Gancedo C, Gancedo JM. 1993. Trehalose-6-phosphate, a new regulator of yeast glycolysis that inhibits hexokinases. *FEBS Lett* 329: 51–4.
- Bloemen HJJ, Wu L, van Gulik WM, Heijnen JJ, Verhaegen MHG. 2003. Reconstruction of the O₂ uptake rate and CO₂ evolution rate on a time scale of seconds. *AIChE J* 49: 1895–908.
- Boles E, Schulte F, Miosga T, Freidel K, Schluter E, Zimmermann FK, Hollenberg CP, Heinisch JJ. 1997. Characterization of a glucose-repressed pyruvate kinase (Pyk2p) in *Saccharomyces cerevisiae* that is catalytically insensitive to fructose-1,6-bisphosphate. *J Bacteriol* 179: 2987–93.
- Brown CJ, Todd KM, Rosenzweig RF. 1998. Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. *Mol Biol Evol* 15: 931–42.
- Brown CM, Hough JS. 1965. Elongation of yeast cells in continuous culture. *Nature* 206: 676–8.
- Cascante M, Llorens M, Melendez-Hevia E, Puigjaner J, Montero F, Marti E. 1996. The metabolic productivity of the cell factory. *J Theor Biol* 182: 317–25.
- de Jong-Gubbels P, Vanrolleghem P, Heijnen S, van Dijken JP, Pronk JT. 1995. Regulation of carbon metabolism in chemostat cultures of *Saccharomyces cerevisiae* grown on mixtures of glucose and ethanol. *Yeast* 11: 407–18.
- Ditzelmuller G, Wohrer W, Kubicek CP, Rohr M. 1983. Nucleotide pools of growing,

- synchronized and stressed cultures of *Saccharomyces cerevisiae*. Arch Microbiol 135: 63–7.
- Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, Botstein D. 2002. Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 99: 16144–9.
- Ferea TL, Botstein D, Brown PO, Rosenzweig RF. 1999. Systematic changes in gene expression patterns following adaptive evolution in yeast. Proc Natl Acad Sci U S A 96: 9721–6.
- Giersch C. 1995. Determining elasticities from multiple measurements of flux rates and metabolite concentrations. application of the multiple modulation method to a reconstituted pathway. Eur J Biochem 227: 194–201.
- Gombert AK, Moreira dos Santos M, Christensen B, Nielsen J. 2001. Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. J Bacteriol 183: 1441–51.
- Haeckel R, Hess B, Lauterborn W, Wuster KH. 1968. Purification and allosteric properties of yeast pyruvate kinase. Hoppe Seylers Z Physiol Chem 349: 699–714.
- Heijnen JJ, Van Gulik WM, Shimizu H, Stephanopoulos G. 2004. Metabolic flux control analysis of branch points: an improved approach to obtain flux control coefficients from large perturbation data. Metab Eng 6: 391–400.
- Hess B, Haeckel R, Brand K. 1966. FDP-activation of yeast pyruvate kinase. Biochem Biophys Res Commun 24: 824–31.
- Kresnowati MT, van Winden WA, Heijnen JJ. 2005. Determination of elasticities, concentration and flux control coefficients from transient metabolite data using linlog kinetics. Metab Eng 7: 142–53.
- Lange HC, Eman M, van Zuijlen G, Visser D, van Dam JC, Frank J, de Mattos MJ, Heijnen JJ. 2001. Improved rapid sampling for *in vivo* kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. Biotechnol Bioeng 75: 406–15.
- Lange HC, Heijnen JJ. 2001. Statistical reconciliation of the elemental and molecular biomass composition of *Saccharomyces cerevisiae*. Biotechnol Bioeng 75: 334–44.
- Lay D. 2003. Linear algebra and its applications. Addison-Wesley, Boston.
- Li Y, Rivera D, Ru W, Gunasekera D, Kemp RG. 1999. Identification of allosteric sites in rabbit phosphofructo-1-kinase. Biochemistry 38: 16407–12.
- Macfarlane N, Ainsworth S. 1972. A kinetic study of baker's-yeast pyruvate kinase activated by fructose 1,6-diphosphate. Biochem J 129: 1035–47.
- Maier A, Volker B, Boles E, Fuhrmann GF. 2002. Characterisation of glucose transport in *Saccharomyces cerevisiae* with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. FEMS Yeast Res 2: 539–50.
- Mashego MR, Jansen ML, Vinke JL, van Gulik WM, Heijnen JJ. 2005. Changes in the metabolome of *Saccharomyces cerevisiae* associated with evolution in aerobic glucose-limited chemostats. FEMS Yeast Res 5: 419–30.
- Mashego MR, van Gulik WM, Vinke JL, Heijnen JJ. 2003. Critical evaluation of sampling techniques for residual glucose determination in carbon-limited chemostat culture of *Saccharomyces cerevisiae*. Biotechnol Bioeng 83: 395–9.
- Murcott TH, Gutfreund H, Muirhead H. 1992. The cooperative binding of fructose-1,6-

- bisphosphate to yeast pyruvate kinase. *Embo J* 11: 3811–4.
- Nissler K, Otto A, Schellenberger W, Hofmann E. 1983. Similarity of activation of yeast phosphofructokinase by AMP and fructose-2,6-bisphosphate. *Biochem Biophys Res Commun* 111: 294–300.
- Novick A, Szilard L. 1950. Description of the chemostat. *Science* 112: 715–6.
- Parmeggiani A, Bowman RH. 1963. Regulation of phosphofructokinase activity by citrate in normal and diabetic muscle. *Biochem Biophys Res Commun* 12: 268–73.
- Portela P, Howell S, Moreno S, Rossi S. 2002. *In vivo* and *in vitro* phosphorylation of two isoforms of yeast pyruvate kinase by protein kinase A. *J Biol Chem* 277: 30477–87.
- Przybylski F, Otto A, Nissler K, Schellenberger W, Hofmann E. 1985. Effects of fructose 1,6-bisphosphate on the activation of yeast phosphofructokinase by fructose 2,6-bisphosphate and AMP. *Biochim Biophys Acta* 831: 350–2.
- Reich JG, Sel'kov EE. 1981. Energy metabolism of the cell. Academic Press, London.
- Rizzi M, Baltés M, Theobald U, Reuss M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol Bioeng* 55: 592–608.
- Salas ML, Vinuela E, Salas M, Sols A. 1965. Citrate inhibition of phosphofructokinase and the pasteur effect. *Biochem Biophys Res Commun* 19: 371–6.
- Shimizu H, Tanaka H, Nakato A, Nagahisa K, Kimura E, Shioya S. 2003. Effects of the changes in enzyme activities on metabolic flux redistribution around the 2-oxoglutarate branch in glutamate production by *Corynebacterium glutamicum*. *Bio-process Biosyst Eng* 25: 291–8.
- Sols A. 1981. Multimodulation of enzyme activity. *Curr Top Cell Regul* 19: 77–101.
- Stephanopoulos G, Aristidou AA, Nielsen J. 1998. Metabolic engineering: principles and methodologies. Academic Press, San Diego.
- Stephanopoulos G, Simpson T. 1997. Flux amplification in complex metabolic networks. *Chem. Eng. Sci.* 52: 2607–27.
- Stuckrath I, Lange HC, Kotter P, van Gulik WM, Entian KD, Heijnen JJ. 2002. Characterization of null mutants of the glyoxylate cycle and gluconeogenic enzymes in *S. cerevisiae* through metabolic network modeling verified by chemostat cultivation. *Biotechnol Bioeng* 77: 61–72.
- Teusink B. 1999. Exposing a complex metabolic system: glycolysis in *Saccharomyces cerevisiae*. PhD Thesis. Department of Microbiology, University of Amsterdam, Amsterdam.
- Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, Snoep JL. 2000. Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem* 267: 5313–29.
- Teusink B, Walsh MC, van Dam K, Westerhoff HV. 1998. The danger of metabolic pathways with turbo design. *Trends Biochem Sci* 23: 162–9.
- Theobald U, Mailinger W, Baltés M, Reuss M, Rizzi M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng* 55: 305–16.
- van Dam JC, Eman M, Frank J, Lange HC, van Dedem GW, Heijnen JJ. 2002. Analysis of glycolytic intermediates in *Saccharomyces cerevisiae* using anion exchange chromatog-

- raphy and electrospray ionization with tandem mass spectrometric detection. *Anal Chim Acta* 460: 209–18.
- van der Heijden RTJM, Heijnen JJ, Hellinga C, Romein B, Luyben KCAM. 1994. Linear constraint relations in biochemical reaction systems. II. Diagnosis and estimation of gross errors. *Biotechnol Bioeng* 43: 11–20.
- van Gulik WM, de Laat WT, Vinke JL, Heijnen JJ. 2000. Application of metabolic flux analysis for the identification of metabolic bottlenecks in the biosynthesis of penicillin-G. *Biotechnol Bioeng* 68: 602–18.
- van Winden WA. 2002. ^{13}C -Labeling technique for metabolic network and flux analysis. Theory and applications. PhD Thesis. Delft University of Technology, Delft.
- van Winden WA, van Dam JC, Ras C, Kleijn RJ, Vinke JL, van Gulik WM, Heijnen JJ. 2005. Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of ^{13}C -labeled primary metabolites. *FEMS Yeast Res* 5: 559–68.
- Vanrolleghem PA, de Jong-Gubbels P, van Gulik WM, Pronk JT, van Dijken JP, Heijnen S. 1996. Validation of a metabolic network for *Saccharomyces cerevisiae* using mixed substrate studies. *Biotechnol Prog* 12: 434–48.
- Verduyn C, Postma E, Scheffers WA, Van Dijken JP. 1992. Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 8: 501–17.
- Visser D, Heijnen JJ. 2002. The mathematics of metabolic control analysis revisited. *Metab Eng* 4: 114–23.
- Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ. 2004a. Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. *Metab Eng* 6: 378–90.
- Visser D, van Zuylen GA, van Dam JC, Eman MR, Proll A, Ras C, Wu L, van Gulik WM, Heijnen JJ. 2004b. Analysis of *in vivo* kinetics of glycolysis in aerobic *Saccharomyces cerevisiae* by application of glucose and ethanol pulses. *Biotechnol Bioeng* 88: 157–67.
- Voet D, Voet JG. 2004. *Biochemistry*. Wiley, Hoboken, 3rd edition.
- Wu L, Lange HC, Van Gulik WM, Heijnen JJ. 2003. Determination of *in vivo* oxygen uptake and carbon dioxide evolution rates from off-gas measurements under highly dynamic conditions. *Biotechnol Bioeng* 81: 448–58.
- Wu L, Wang W, van Gulik WM, Heijnen JJ. 2004. A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. *Eur J Biochem* 271: 3348–59.
- Yoshino M, Murakami K. 1982. AMP deaminase as a control system of glycolysis in yeast. Mechanism of the inhibition of glycolysis by fatty acid and citrate. *J Biol Chem* 257: 10644–9.

Appendix

The effector concentrations in glycolysis (see Table 6.4) were obtained from Mashego et al. (2005) for 13 steady states during prolonged chemostat PB. The dimension of the resulting matrix \mathbf{A} is 13×3 , 13×4 and 13×3 for the irreversible reaction catalyzed by HK, PFK and PK respectively. According to Eq. 6.5, the elasticity

vector $\mathbf{E}^{\mathbf{x}0}$ should lie in the null space of \mathbf{A} . Due to measurement inaccuracies the null space is however empty. To obtain linear constraints of the elasticities, the matrix \mathbf{A} therefore has to be decomposed (Lay, 2003):

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{A1}$$

where the matrix \mathbf{S} contains singular values of \mathbf{A} . Table A1 shows that for each irreversible reaction in glycolysis, two singular values of \mathbf{A} obtained from the singular value decomposition are significantly closer to zero than others. These singular values were set to 0 in \mathbf{S} , with which the matrix \mathbf{A} is recalculated according to Eq. A1. Now the null space of the augmented matrix \mathbf{A} is spanned by a number of independent vectors equaling the number of singular values that were set at zero. Hence, $\mathbf{E}^{\mathbf{x}0}$ is expressed by a linear combination of two linearly independent vectors \mathbf{v}_1 and \mathbf{v}_2 :

$$\varepsilon^{\mathbf{x}0} = [\mathbf{v}_1 \quad \mathbf{v}_2] \begin{bmatrix} a \\ b \end{bmatrix} \tag{A2}$$

These linearly independent vectors obtained from multiple steady state measurements are given in Table A2. Eliminating the scalars a and b with two freely chosen elasticities, Eq. A2 can be easily rewritten into an algebraic relationship between the individual elasticities, shown in Table A2.

Table A1 Singular values of matrix \mathbf{A} associated with the irreversible reactions in glycolysis

| HK | PFK | PK |
|------|------|------|
| 4.94 | 5.62 | 5.97 |
| 0.70 | 1.28 | 1.35 |
| 0.34 | 0.20 | 0.47 |
| | 0.15 | |

Table A2 Null space of matrix \mathbf{A} associated with the irreversible reactions in glycolysis

| HK | \mathbf{v}_1 | \mathbf{v}_2 | PFK | \mathbf{v}_1 | \mathbf{v}_2 | PK | \mathbf{v}_2 | \mathbf{v}_2 |
|-------------------|----------------|----------------|--------|----------------|----------------|--------|----------------|----------------|
| Glc ^{Ex} | 1.00 | 1.00 | F6P | 1.00 | 1.00 | PEP | 1.00 | 1.00 |
| G6P | -0.99 | 0.41 | F1,6bP | -1.33 | -0.17 | F1,6bP | -0.57 | 19.37 |
| T6P | 0.49 | -1.21 | F2,6bP | 1.71 | -0.98 | Cit | -0.97 | -10.40 |
| | | | Cit | -1.26 | -0.35 | | | |

Chapter 7

In Vivo Kinetic Analysis by Carbon, Electron and ATP Balances

Abstract Models of *in vivo* enzyme kinetics are important in understanding metabolic control mechanisms in microorganisms and facilitate rational strain improvement by metabolic engineering. This paper investigates *in vivo* kinetics of the primary metabolism in *S. cerevisiae* by applying a glucose pulse to an aerobic carbon-limited chemostat culture, followed by comprehensive metabolome analysis over a subsequent transient of 300 seconds. The metabolome responses reveal three distinct metabolic phases during the transient and in each phase, cumulative carbon, electron and ATP balances were constructed to identify significant carbon, electron and energy sinks. In phase I (0 – 50 seconds after the pulse), the glucose uptake rate (q_{Glc}) increased immediately to 17 fold the steady state level q_{Glc}^0 . About 2% of consumed glucose was converted to ethanol in phase I, while 60% accumulated as glycolytic and storage metabolites, which caused an energy depletion manifested in a drop in the ATP level. This likely led to the subsequent drop of q_{Glc} to 3 fold q_{Glc}^0 . In phase II (50 – 150 seconds), q_{Glc} increased gradually from 3 to 8 fold q_{Glc}^0 . Fermentative metabolism set in and became the major carbon and electron sink. In phase III (150 – 300 seconds), q_{Glc} remained at 8 fold q_{Glc}^0 and a quasi steady state was established. 29% of the consumed glucose was not recovered by metabolome measurements and might be partially channeled into anabolic processes. This is supported by macroscopic balances of extracellular fluxes and ATP. The *in vivo* kinetic mechanisms, flux distribution and equilibrium relationships underlying the observed responses are discussed.

7.1 Introduction

Knowledge of *in vivo* enzyme kinetics in microorganisms is important in understanding metabolic control mechanisms operating on the level of the metabolome and can be used to assist rational redesign of metabolic pathways to enhance desired functionalities of microbes (Wiechert, 2002). *In vivo* kinetics can be studied by stimulus-response experiments, in which cells grown in a (quasi) steady state are perturbed by an external stimulus and the dynamic responses of e.g. the intra- and extracellular metabolites are monitored. The time window of observation is usually within tens to a few hundred seconds, such that responses of the metabolome can be mainly attributed to rapid (allosteric) enzyme-metabolite interactions (Theobald et al., 1997). *In vivo* kinetic parameters can be estimated from such measured responses, based on a set of mass balances for each metabolite species (Chassagnole et al., 2002; Rizzi et al., 1997; Vaseghi et al., 1999). The stimulus-response methodology is an ideal tool to obtain rich kinetic information and has been applied to various microorganisms under different growth conditions (Chassagnole et al., 2002; Ostergaard et al., 2001; Schmitz et al., 2002; Theobald et al., 1997; Visser et al., 2004).

For *S. cerevisiae*, a frequently applied perturbation is the sudden transition from sugar limitation to sugar excess, realized by e.g. the instantaneous addition of a concentrated glucose solution (i.e. a glucose pulse) to a glucose-limited chemostat culture. This leads to a sudden increase of the glucose influx and the secretion of ethanol, acetate and glycerol. In the cell, large concentration changes have been observed for intermediates and cofactors involved in glycolysis (Theobald et al., 1997). These measured transients were combined in a rigorous kinetic model focusing on glycolysis (Rizzi et al., 1997), which was extended later to include the pentose phosphate (PP) pathway (Vaseghi et al., 1999).

Characterization of the transient state after the perturbation is however still far from complete. Limited information has been obtained for other metabolic pathways, such as the TCA cycle and anabolic processes, which have been modeled so far as lumped reactions with simplified kinetics. These pathways might however lead to unknown carbon sinks, which can potentially bias parameter estimates if they were to make up a significant fraction of the total uptake of the added glucose. Moreover, carbon and energy metabolism is highly connected in *S. cerevisiae*. Excess glucose is distributed among carbon sinks related to catabolic, energy providing processes as well as anabolic, energy consuming processes (such as storage). Knowledge of carbon sinks therefore provides insight into the interaction of carbon and energy metabolism during the transient.

Carbon and electron balances can be employed to evaluate whether all significant carbon sinks and the related metabolic processes are accounted for in a metabolic model. In a (quasi) steady state, these balances are prerequisite for adequate metabolic flux analysis (Stephanopoulos et al., 1998). During a transient, dynamic carbon balances have been constructed in a time window of tens of minutes to hours (Herwig et al., 2001; Ostergaard et al., 2001; van Urk et al., 1988), based on measured extracellular fluxes and an assumed intracellular quasi steady state. It has however not been attempted so far to set up complete carbon, electron and ATP balances within the time scale of interest for stimulus response experiments, namely

tens to hundreds of seconds.

To obtain a more comprehensive characterization of the dynamic responses in both catabolic and anabolic pathways, we performed an in-depth analysis of the yeast metabolome over a 300-seconds transient after the addition of a glucose pulse to an aerobic carbon-limited chemostat culture of *S. cerevisiae*. In addition to glycolytic intermediates and adenine nucleotides, we quantified intracellular free amino acids, intermediates in the TCA cycle, pentose phosphate pathway and storage carbon metabolism, as well as the *in vivo* O₂ uptake rate and CO₂ evolution rate. These measurements allow the construction of cumulative carbon, electron and ATP balances, as well as a balance of intracellular phosphate, in different metabolic regimes during the transient, from which significant carbon, electron and ATP sources and sinks were identified. In addition, the metabolome measurements were applied to obtain insight into *in vivo* equilibria, flux distribution and kinetic mechanisms, which are needed for the further construction of an *in vivo* kinetic model of the primary metabolism of *S. cerevisiae*.

7.2 Materials and methods

Yeast strain and chemostat cultivation The haploid *S. cerevisiae* strain CEN.-PK 113-7D was cultivated in an aerobic, carbon-limited chemostat with a working volume of 4 l at a dilution rate of 0.051 h⁻¹, as described previously (Lange et al., 2001). The medium used for chemostat cultivation was based on Lange et al. (2001) with 27.1 g/l glucose and 1.42 g/l ethanol to obtain a biomass concentration of approximately 15 g_{DW}/l in the steady state. The dissolved oxygen tension (DOT) in the fermentor and the O₂ and CO₂ content in the fermentor off-gas were analyzed as described previously (Wu et al., 2003). The DOT was maintained above 70% in the steady state.

Rapid sampling, quenching and metabolite extraction Culture supernatant for the analysis of extracellular metabolites was obtained by rapid quenching of the broth with stainless steel beads at -20°C followed by filtration, adapted from Mashego et al. (2003) for a sample volume of 2 ml. For the analysis of intracellular metabolites, sampling and sample preparation was carried out as described previously (Mashego et al., 2004). Briefly, approximately 1 ml of broth was rapidly quenched in 5 ml of -40°C 60% (v/v) aqueous methanol solution. After centrifugation and decanting of the liquid, the biomass pellet was resuspended in 5 ml of -40°C 60% (v/v) aqueous methanol solution and again centrifuged. These steps were carried out below -20°C. Intracellular metabolites were extracted from the biomass pellet with boiling 75% (v/v) aqueous ethanol solution. The ethanol extracts were evaporated to dryness. The dried samples was resuspended in 0.5 ml Milli-Q water and centrifuged. The cell extract and insoluble residue was separated and stored at -80°C until analysis.

Determination of culture dry weight Culture samples (5 ml, in triplicate) were filtered through pre-dried and pre-weighed nitrocellulose filters (pore size 0.45 µm, Gelman Science, Ann Arbor, MI, USA). The filters were washed twice with 5 ml demineralized water, dried at 70°C for 48 hours and weighed.

Extracellular metabolite analysis The concentration of glucose, acetate and glycerol in the supernatant was analyzed with Enzytec™ kits (Scil Diagnostics GmbH, Martinsried, Germany). The concentration of ethanol in the supernatant was analyzed with the Enzymatic Bioanalysis kit (R-Biopharm AG, Darmstadt, Germany). Absorbance was read on an Agilent 8453-UV-visible spectroscopy system (Agilent Technologies Deutschland GmbH, Waldbronn, Germany).

MS analysis of intracellular metabolites The concentrations of G6P, F6P, F1,6bP, 2PG, 3PG and PEP in the cell extract were analyzed by liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) (van Dam et al., 2002). The same analysis method and conditions were applied for the analysis of Pyr, Cit, iCit, α KG, Suc, Fum, Mal, G1P, M6P, T6P, F2,6bP and Pi in the cell extract. The metabolites 2PG and 3PG could not be resolved with the applied analytical procedure, the same held for Cit and iCit. Consequently, only the sum of these compounds was determined. For F2,6bP, only the peak areas were quantified instead of the absolute amount in the sample.

The concentrations of the adenine nucleotides (AMP, ADP, ATP) in the cell extract were quantified by LC-ESI-MS/MS. The nucleotides were separated by an ion pairing reversed phase HPLC method adapted from (Claire, 2000), using a XTerra MS C18 column (100 mm \times 1 mm) equipped with a guard column (10 mm \times 2.1 mm) (both from Waters, Milford, USA). The standard solution and sample injection volume was 10 μ l; all standard solutions and samples were mixed 24:1 with a 2M solution of the ion pairing reagents tetrabutylammoniumacetate (TBAA, Alrich, Steinheim, Germany) to obtain a final TBAA concentration of 80mM. The chromatography was performed at room temperature with an Alliance HT 2795 pump system (Waters, Milford, USA), giving an isocratic elution flow of 0.1 ml/min. The mobile phase was a 10 mM $\text{NH}_4\text{H}_2\text{PO}_4$ solution that was adjusted to pH 6.4 with NH_4OH , after which 2 mM tetrabutylammoniumhydroxide (TBAH, Aldrich, Steinheim, Germany) was added, giving a pH of 6.8, followed by addition of 15% (v/v) acetonitril solution. The post-column eluent was mixed with 0.1 ml/min 80% (v/v) acetonitril solution. The solutions were filtered (0.45 μ m, Gelman Sciences, Michigan, USA) prior to use to remove particles and bubbles. The HPLC was coupled to the MS/MS (Quattro Ultima Pt, Micromass Ltd., UK) via a flow splitter which sent a flow of 0.06 ml/min to ESI that operated in positive mode with a nebulizer gas (nitrogen) flow of 50 l/h and a desolvation gas (nitrogen) flow of 500 l/h and 300°C. The source block temperature was 120°C. The capillary voltage was set at 3.5 kV and the cone voltage at 35 V. The $[\text{M} + \text{H}]^+$ ions (m/z of 348, 428 and 508 for respectively AMP, ADP and ATP) were fragmented by collisionally induced dissociation with argon (collision energy 24 V). The strongest daughter fragment of m/z 136 was monitored for each of them.

Intracellular trehalose analysis Intracellular trehalose was analyzed by quantitative ^1H -NMR. Analysis was performed at 360 MHz on a Bruker AMX 360 spectrometer (Bruker Analytik, Karlsruhe, Germany). To 0.5 ml of cell extract, an equal amount of a standard solution containing maleic acid and EDTA was added. After lyophilization, the residue was dissolved in D_2O and the ^1H NMR spectrum was measured, using a relaxation delay of 30 seconds, ensuring full relaxation of all the hydrogen atoms between pulses. The integrals of the ^1H protons of trehalose (dou-

blet at 5.24 ppm) and the internal standard (singlet at 6.1 ppm) were measured and the content of trehalose was calculated.

Intracellular glycogen analysis Intracellular glycogen was determined in the cell extracts and in the insoluble residues, as described in Parrou and Francois (1997). Glycogen in the cell extracts was directly hydrolyzed with amyloglucosidase without alkaline extraction.

Intracellular free amino acid analysis Intracellular concentrations of free amino acids were measured by HPLC using the AccQ-tag system (Vriezen et al., 1997).

Estimation of extracellular Pi carryover to cell extracts About 0.25 ml out of the 5 ml quenching or washing solution remained in the test tube after decanting, which corresponds to a washing efficiency of $1 - (0.25/5)^2 = 99.75\%$ and a carryover of 0.25% after the quenching and the washing step. A steady state extracellular Pi concentration of 39.3 mM was calculated based on the biomass phosphate content (Lange and Heijnen, 2001) and the medium composition. Assuming constant extracellular Pi concentration during the transient, the carryover is no more than 8% of the measured Pi concentration in the cell extract.

Glucose pulse experiment When *S. cerevisiae* was grown to about 10 generations in an aerobic carbon-limited chemostat, 15 ml of a glucose solution in a syringe (266.7 g/l, thus containing 4 g of glucose) was injected into the fermentor by a pneumatic device. Samples for the extra- and intracellular metabolite analysis were taken both during the steady state shortly before the glucose pulse and during a 300-seconds transient after the perturbation. During the transient, the feed to the fermentor was continued but no effluent was removed via the effluent pump. The latter was due to the fact that the chemostat was weight-controlled and the fermentor weight decreased slightly due to sampling.

Steady state metabolic flux analysis The extracellular fluxes in the steady state were reconciliated and used to estimate intracellular fluxes, as described in Chapter 6 of this thesis.

Estimation of *in vivo* O₂ uptake rate and CO₂ evolution rate The *in vivo* O₂ uptake rate and CO₂ evolution rates (q_{O_2} and q_{CO_2}) during the transient were reconstructed from the measured concentrations of O₂ and CO₂ in the off-gas and the measured DOT, using a dynamic mass transfer model (Bloemen et al., 2003). The cumulative O₂ uptake and CO₂ evolution was obtained by numerically integrating the obtained q_{O_2} and q_{CO_2} .

Construction of cumulative carbon, electron and ATP balances The cumulative carbon balance in Cmol per Cmol_{biomass} (Cmol/Cmol_X compares the cumulative uptake of carbon sources against the accumulation of known carbon sinks, calculated at each sampling time point relative to t_0 , at which the glucose pulse was given (see also Herwig et al. 2001). During the transient, glucose (added both via the pulse and the feed) is the sole carbon source, assuming that the small amount of ethanol in the feed is not metabolized. The known carbon sinks comprise of measured secreted extracellular metabolites, measured intracellular metabolites, CO₂

and biomass. The biomass growth rate during the transient was assumed the same as in the steady state, i.e. 0.05 h^{-1} . The adenine nucleotides were excluded from the carbon balances, assuming that the adenine moiety is conserved during the short time window. A constant total broth volume was assumed. This is allowed since within 300 seconds, effluent via sampling ($\sim 0.1 \text{ l}$) and influent via the feed pump (0.02 l) is negligible compared to the total broth volume (4 l). These assumptions were also applied in the cumulative electron and ATP balances.

The cumulative electron balance (in $\text{mol}_e/\text{Cmol}_X$) was set up in a similar way as the cumulative carbon balance, but now using degree of reduction of the metabolites, calculated according to Nielsen and Villadsen (1994). It compares the cumulative uptake of electron sources against the accumulation of known electron sinks. During the transient, glucose is the sole electron source. The known electron sinks include the considered carbon sinks and O_2 .

The cumulative ATP balance (in $\text{mol ATP equivalents}/\text{Cmol}_X$) was set up in a similar way as the cumulative carbon balance. It compares the cumulative consumption with the cumulative production of ATP equivalents. Table 7.1 summarizes the amount of consumption and production of ATP equivalents associated with the carbon sinks considered in the ATP balance. The intermediates in the PP pathway and the TCA cycle, as well as free amino acids were excluded from the ATP balance as they account only for a negligible share of the total carbon uptake (see Results and Discussion and Table 7.4).

Table 7.1 Energy production or consumption associated with the accumulation of metabolites

| Metabolite | $\text{mol}_{\text{ATPeq}}/\text{mol}$ | Metabolite | $\text{mol}_{\text{ATPeq}}/\text{mol}$ |
|-----------------|--|------------|--|
| ATP production | | | |
| Ethanol | 1 | | |
| Acetate | 1 ^a | | |
| O_2 | 2 ^b | | |
| ATP consumption | | | |
| Glucose | 0 ^a | Glycerol | 1 |
| Biomass | 1.65 ^b | G6P | 1 |
| F6P | 1 | F1,6bP | 2 |
| 2PG+3PG | 0 | PEP | 0 |
| G1P | 1 | T6P | 3 |
| Trehalose | 3 | M6P | 1 |

^a Assume no ATP cost for the uptake of glucose and secretion of acetate.

^b According to Verduyn et al. (1991). Effective P/O ratio is assumed to be 1.

7.3 Results and discussion

Characteristic responses of extracellular fluxes and intracellular metabolites to a glucose pulse Prior to the glucose pulse, the culture of *S. cerevisiae* was maintained in a reference steady state at a growth rate of 0.05 h^{-1} . The measured metabolite concentrations in the steady state are summarized in Table 7.2. The reconciliated extracellular fluxes and the estimated intracellular fluxes are given in Table 7.3 and Fig 7.1, respectively.

This reference steady state was perturbed by a glucose pulse, increasing the residual glucose concentration from $27.7 \pm 0.6 \text{ mg/l}$ by 1.0 g/l . Fig. 7.2 shows that the uptake of excess glucose led to the secretion of ethanol, acetate and a small amount of glycerol. The biomass specific fluxes of glucose, ethanol and acetate (q_{Glc} , q_{EtOH} and q_{Act}) were derived from the measured extracellular concentration profiles. These were shown in Fig. 7.3, together with the estimated biomass specific q_{O_2} , q_{CO_2} and the respiratory quotient (RQ). Responses of intracellular metabolites are summarized in Fig. 7.4-7.7, in terms of fold changes relative to the concentrations in the reference steady state. From the flux and metabolome responses, three sequential metabolic phases can be discerned during the 300-seconds transient. The phases are indicated by vertical dotted lines in the figures.

Table 7.2 Steady state concentrations of different classes of intracellular metabolites (in $\mu\text{mol/Cmol}_X$)

| Metabolite | Concentration | Metabolite | Concentration |
|---------------------------|-----------------------|---------------------|---------------|
| Glycolytic intermediates | | Amino acids | |
| G6P | 53 ± 2.6 | Asp | 36 ± 0.4 |
| F6P | 9.9 ± 0.8 | Ser | 8.3 ± 2.0 |
| F1,6bP | 7.1 ± 0.7 | Asn | 30 ± 1.4 |
| 2PG+3PG | 24 ± 1.2 | Glm | 362 ± 4.4 |
| PEP | 25 ± 2.4 | Gln | 142 ± 1.6 |
| Pyr | 2.2 ± 0.2 | His | 36 ± 1.6 |
| | | Thr | 11 ± 0.1 |
| TCA cycle intermediates | | Ala | 55 ± 0.7 |
| Cit + iCit | 140 ± 5.1 | Arg | 47 ± 0.7 |
| α KG | 2.2 ± 0.2 | Pro | 7.9 ± 1.4 |
| Suc | 1.5 ± 0.3 | Tyr | 6.8 ± 0.2 |
| Fum | 1.4 ± 0.2 | Val | 20 ± 0.7 |
| Mal | 7.3 ± 0.8 | Orn | 27 ± 0.7 |
| | | Lys | 23 ± 0.3 |
| PP pathway intermediates | | Ile | 5.7 ± 0.1 |
| 6PG | 6.1 ± 0.6 | Leu | 35 ± 0.7 |
| | | Phe | 3.2 ± 0.3 |
| Storage carbon metabolism | | Adenine Nucleotides | |
| G1P | 7.8 ± 2.2 | ATP | 192 ± 0.3 |
| T6P | 18 ± 0.9 | ADP | 45 ± 1.3 |
| Trehalose | $2 \cdot 10^3 \pm 38$ | AMP | 12 ± 1.5 |
| M6P | 22 ± 4.2 | | |

Table 7.3 Measured biomass specific extracellular fluxes in the steady state

| | mol/Cmol _X /h | | |
|------------|--------------------------|------------|--------|
| q_{Glc} | -0.013 | q_{EtOH} | -0.003 |
| q_{Act} | 0.000 | q_{Glyc} | 0.000 |
| q_{CO_2} | 0.037 | q_{O_2} | -0.038 |
| q_X | 0.051 | | |

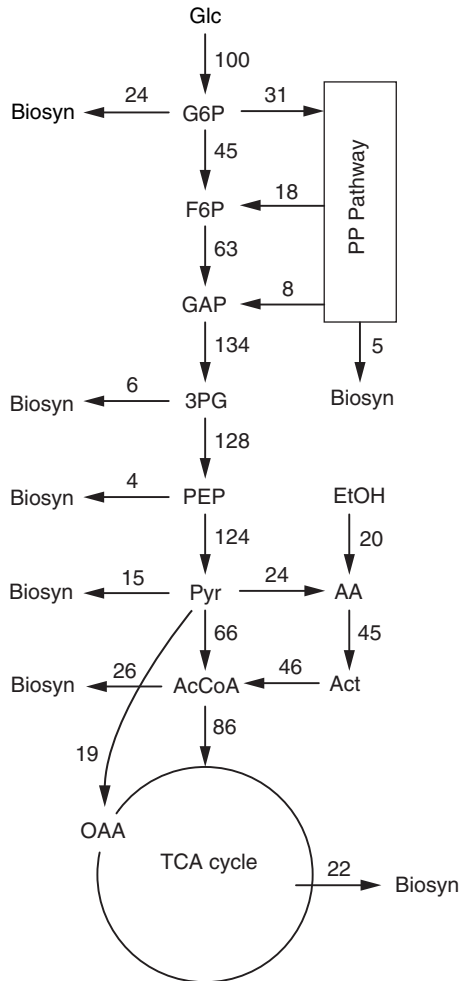


Figure 7.1 Estimated intracellular fluxes in the reference steady state (normalized against a glucose influx of 0.013 mol/Cmol_X/h)

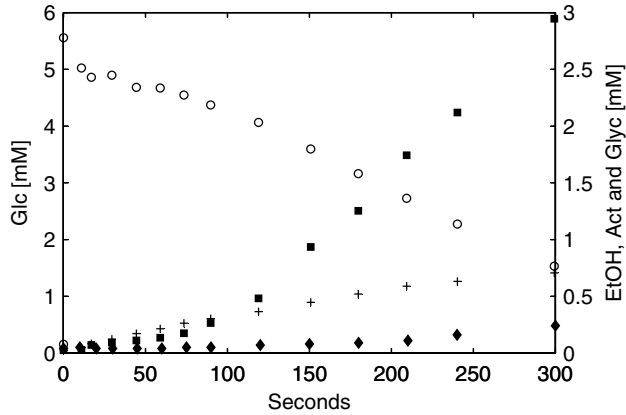


Figure 7.2 Concentration profiles of extracellular metabolites after a glucose pulse added at $t = 0$. Glucose: o, ethanol: ■, acetate: +, glycerol: ◆.

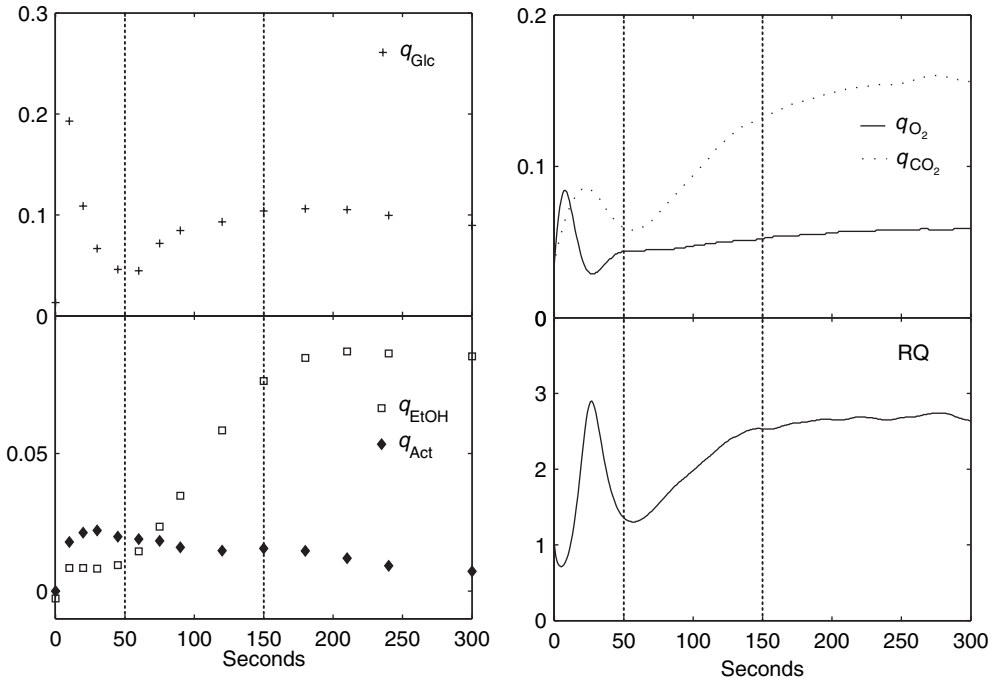


Figure 7.3 Estimated biomass specific fluxes of extracellular metabolites in $\text{mmol/Cmol}_X/\text{h}$.

Highly dynamic changes occurred to extracellular fluxes and intracellular metabolite concentrations in phase I (0 – 50 seconds after the glucose pulse). The q_{Glc} increased to 17 fold the steady state flux (q_{Glc}^0) at 10 seconds, followed by a sharp drop to about 3 times q_{Glc}^0 at 50 seconds. Similar dynamic patterns are found for q_{O_2} and q_{CO_2} . Surprisingly, ethanol production in this phase was not significant and much

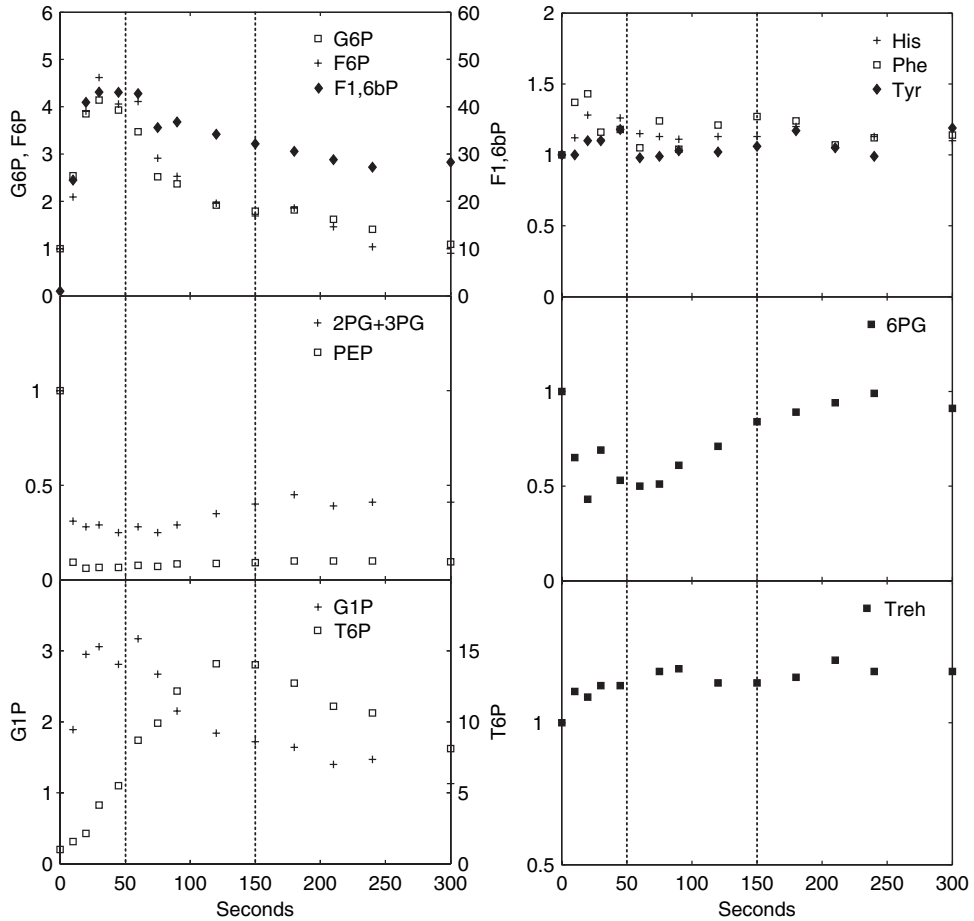


Figure 7.4 Dynamic responses of intermediates in glycolysis, the PP pathway and storage carbon metabolism and thence derived biosynthetic precursors, in fold changes relative to the reference steady state.

lower than q_{Act} . Intracellular metabolite responses in phase I are characterized by accumulation of the hexose phosphate pool, T6P and trehalose (relative to an already high steady state level, see Table 7.2), decreases in the phosphorylated C3-pool (i.e. 2PG, 3PG and PEP) and 6PG (Fig. 7.4). In addition, peaks in the concentration of pyruvate and several TCA cycle intermediates were observed (Fig. 7.5). The ATP concentration and the energy charge, defined by

$$(ATP + \frac{1}{2}ADP) / (ATP + ADP + AMP)$$

sharply decreased in phase I (Fig. 7.6). Noticeably, the ADP and AMP concentration decreased as well after an initial increase. This leads to a net decrease of the sum of adenine nucleotides, which has also been observed by Theobald et al. (1997).

In phase II (50–150 seconds after the pulse), gradual changes in the extracellular fluxes and intracellular metabolite concentrations were observed. The q_{Glc} increased from 3 to 8 fold q_{Glc}^0 during phase II. The fermentative metabolism set in, evidenced

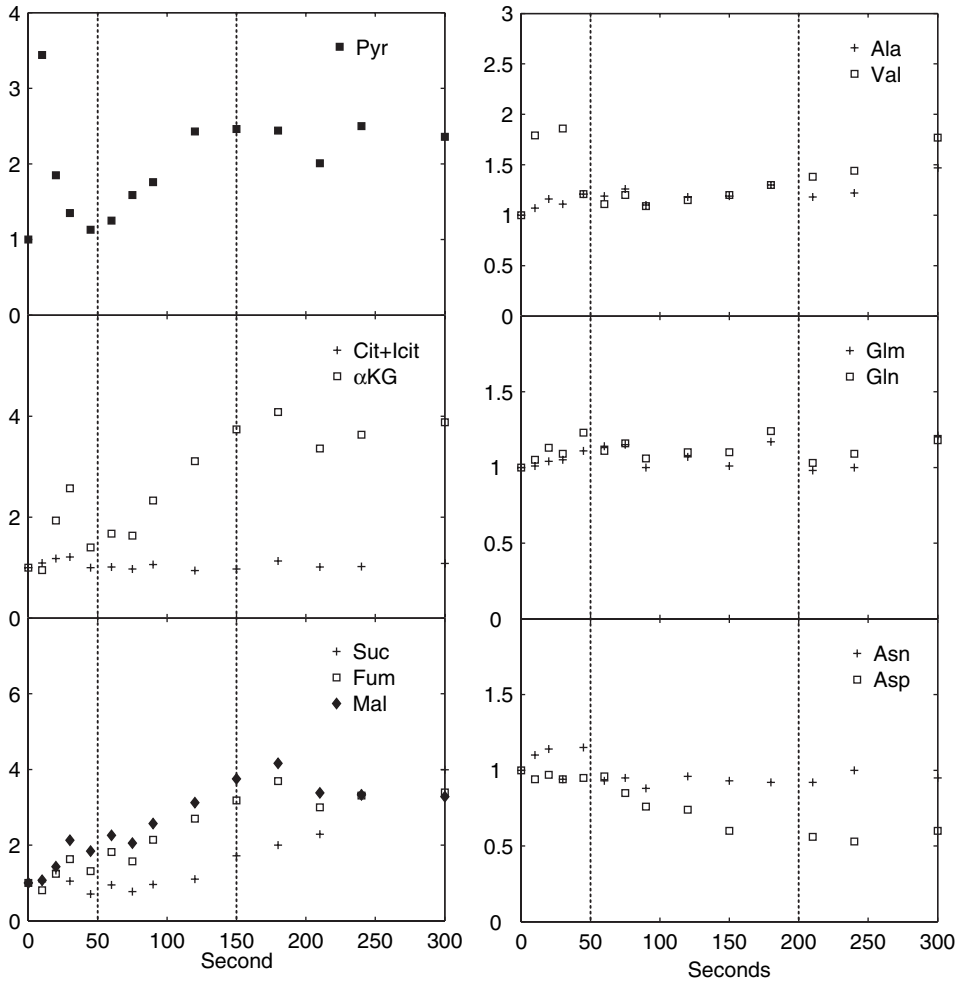


Figure 7.5 Dynamic responses of pyruvate, TCA cycle intermediates and thence derived amino acids in fold changes relative to the reference steady state.

by strong increases in q_{EtOH} , q_{CO_2} and RQ. The ATP level and the energy charge recovered, due to increased ATP generation in both fermentative metabolism (i.e. increased q_{EtOH}) and oxidative phosphorylation (i.e. increased q_{O_2}). In the cell, the concentrations of the hexose phosphates decreased, whereas those of the TCA cycle intermediates, T6P and F2,6bP (Fig. 7.7) increased gradually.

In phase III (150 – 300 seconds after the pulse), the extracellular fluxes and metabolite concentrations remained rather constant. The magnitude of metabolite concentration changes (i.e. $|dx/dt|$) are negligible compared to the magnitude of extracellular fluxes, indicating that an intracellular quasi steady state was established.

Remarkably, no significant changes were observed in intracellular amino acid concentrations in all 3 phases (Fig. 7.4 and 7.5). The only exception of aspartate, with a 2 fold decrease in phase III compared to the reference steady state.

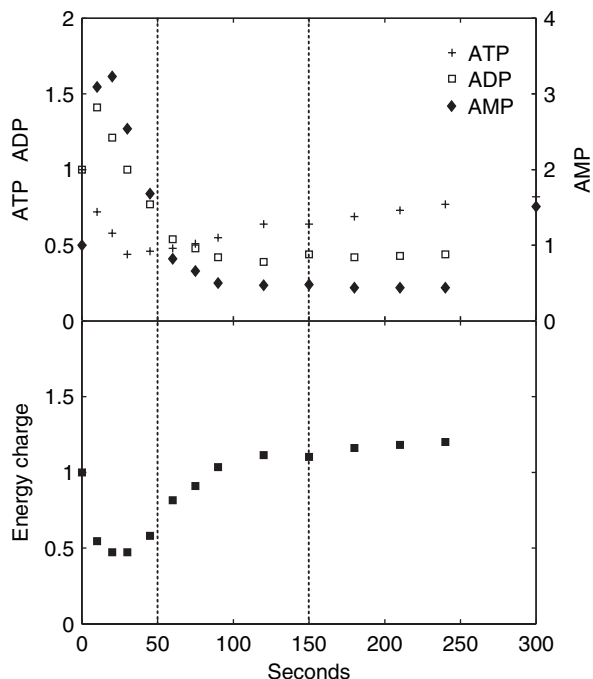


Figure 7.6 Dynamic responses of ATP, ADP, AMP and the energy charge in fold changes relative to the reference steady state. The energy charge in the steady state is 0.86, calculated from the steady state ATP, ADP and AMP concentrations in Table 7.3.

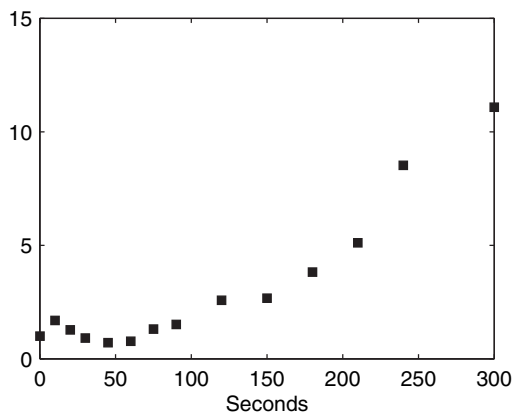


Figure 7.7 Dynamic response of F2,6bP (in fold changes relative to the reference steady state).

Cumulative carbon, electron and ATP balances The available intra- and extracellular metabolite measurements were combined with the estimated highly dynamic *in vivo* q_{O_2} and q_{CO_2} to construct the cumulative carbon, electron and ATP balances, shown in Table 7.4, 7.5 and 7.6. The measured glycogen concentrations in

the insoluble residues and the cell extracts are in the order of 20 and $3.4 \mu\text{mol}/\text{g}_{\text{DW}}$ respectively, much lower than the glycogen content of $650 \mu\text{mol}/\text{g}_{\text{DW}}$ reported for *S. cerevisiae* grown at 0.05 h^{-1} (Guillou et al., 2004). The alkaline extraction seems to be incomplete in our hands, presumably due to the use of insoluble residue obtained after the boiling ethanol treatment, rather than directly using yeast cells as described in Parrou and Francois (1997). Hence, the glycogen measurements were considered unreliable and excluded from the balance calculations.

The cumulative carbon, electron and ATP balances (Fig. 7.8) are evaluated below in each of the three metabolic regimes during the transient.

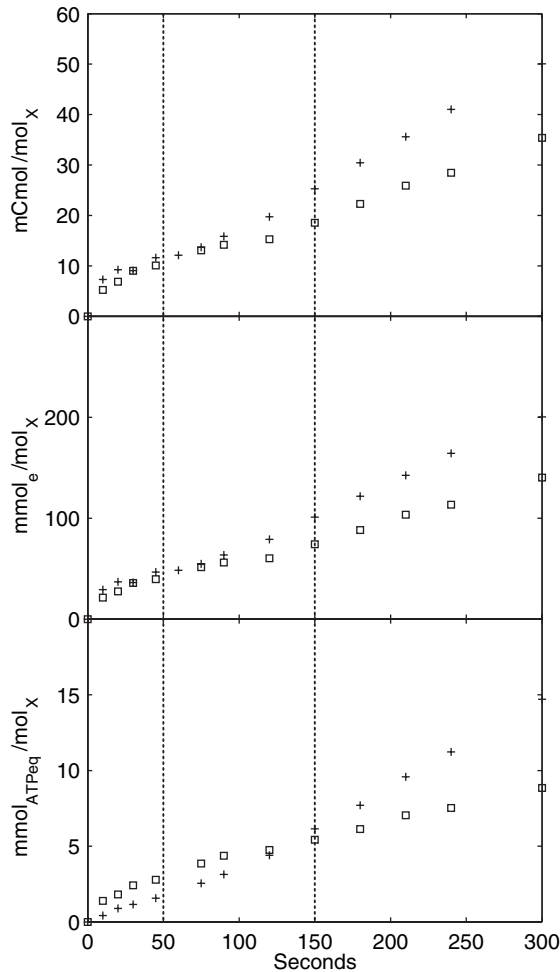


Figure 7.8 Cumulative carbon, electron and ATP balances during the transient. Carbon/electron source and ATP production: +, carbon/electron sink and ATP consumption: □.

Table 7.4 Contribution of different carbon sinks to the cumulative glucose uptake relative to $t = 0$ (in %Cmol/Cmol_{Glc})

| | 45 sec | 150 sec | 300 sec |
|---|--------|---------|---------|
| Cumulative glucose uptake (mCmol/Cmol _X) | 12 | 25 | 50 |
| Carbon sinks | | | |
| Extracellular sinks | | | |
| Ethanol | 2.0 | 12 | 19 |
| Acetate | 4.3 | 5.7 | 4.7 |
| CO ₂ | 7.8 | 14 | 20 |
| Glycerol | 0.0 | 0.8 | 2.0 |
| Biomass ^a | 5.4 | 8.3 | 8.3 |
| Intracellular sinks | | | |
| Glycolytic intermediates | 27 | 6.7 | 2.4 |
| Storage carbon | 36 | 25 | 12 |
| TCA cycle intermediates | 0.1 | 0.4 | 0.4 |
| Amino acids | 5.4 | 0.8 | 1.7 |
| Unknown sinks | 13 | 27 | 29 |

^a A constant growth rate of 0.05 h⁻¹ is assumed.

Table 7.5 Contribution of different electron sinks to the cumulative electron uptake relative to $t = 0$ (in % mol_e/mol_e)

| | 45 sec | 150 sec | 300 sec |
|---|--------|---------|---------|
| Cumulative electron uptake (mmol _e /Cmol _X) | 47 | 101 | 200 |
| Electron sinks | | | |
| Extracellular sinks | | | |
| Ethanol | 3.0 | 17 | 29 |
| Acetate | 4.3 | 5.7 | 4.7 |
| Glycerol | 0.0 | 0.9 | 2.4 |
| Biomass ^a | 5.7 | 8.7 | 8.8 |
| O ₂ | 5.2 | 7.8 | 8.7 |
| Intracellular sinks | | | |
| Glycolytic intermediates | 27 | 6.7 | 2.4 |
| Storage carbon | 37 | 25 | 12 |
| TCA cycle intermediates | 0.2 | 0.3 | 0.3 |
| Amino acids | 5.0 | 0.8 | 1.7 |
| Unknown sinks | 15 | 27 | 30 |

^a A constant growth rate of 0.05 h⁻¹ is assumed.

Table 7.6 Cumulative energy production and consumption relative to $t = 0$

| mmol _{ATP_{eq}} /Cmol _X | | 45 sec | 150 sec | 300 sec |
|---|---|--------|---------|---------|
| Energy generation | Acetate secretion | 0.3 | 0.7 | 1.2 |
| | Ethanol secretion | 0.1 | 1.5 | 4.8 |
| | Oxidative phosphorylation | 1.2 | 4.0 | 8.7 |
| Total | | 1.6 | 6.2 | 14.7 |
| Energy consumption | Biomass ^a | 1.0 | 3.3 | 6.6 |
| | Glycerol secretion | 0.0 | 0.1 | 0.4 |
| | Accumulation of phosphorylated compounds and storage material | 1.8 | 2.1 | 1.8 |
| | Total | 2.8 | 5.4 | 8.9 |

^a A constant growth rate of 0.05 h^{-1} is assumed.

In phase I (0 – 50 seconds), the majority of the measured carbon and electron uptake can be successfully accounted for by the known carbon and electron sinks. Table 7.4 shows the relative contribution of various carbon sinks to the total glucose uptake at the end of each metabolic regime. In phase I, more than 60% of the consumed glucose was directed to (phosphorylated) intracellular metabolites in upper glycolysis and storage carbon metabolism, while 20% was converted to secreted compounds (ethanol, acetate, CO₂ and biomass). The ATP balance (Table 7.6) shows that the ATP consumption exceeded its production, which agrees with the sharp drop of the ATP level and the energy charge. However, at the end of phase I the resulting energy depletion (1.2 mmol ATP_{eq}/Cmol_X) largely exceeded the observed total decrease of ATP (0.1 mmol ATP/Cmol_X). As biomass represents a major fraction in the ATP consumption (Table 7.6) but only a minor fraction in the carbon balance (Table 7.4), this might indicate a temporal cessation of growth in phase I. Alternatively, the discrepancy might be due to yet unaccounted energy generation processes or an underestimated effective P/O ratio.

In summary, in phase I intense dynamics in glycolysis led to elevated pools of phosphorylated compounds and storage material, thereby creating a large energy drain. It is worth noticing that only 2% of the consumed glucose was converted to ethanol. An interesting implication of this finding is that in biomass-directed industrial applications of *S. cerevisiae*, undesired ethanol formation due to inhomogeneous distribution and mixing in large-scale fermentors (Reuss and Bajpai, 1991) might be partly avoided if the mixing time (e.g. the maximum time span that yeast cells encounter a local glucose excess) is reduced to below 50 seconds.

In phase II (50 – 150 seconds), a gap between the carbon/electron sources and sinks emerged at about 100 seconds. At 150 seconds, 26% of the consumed glucose was channeled into unknown sinks. Another 26% was directed to fermentative metabolism, which was the main carbon and electron sink at the end of phase II. The onset of fermentative metabolism and an increased q_{O_2} led to accelerated ATP generation. The ATP balance shows a slight ATP-excess at 150 seconds, which corresponds to the measured increase in the ATP concentration and the recovery of the energy charge.

Table 7.7 Comparison of the carbon, electron and ATP balances in phase III under the following assumption: A. increased growth at 0.14 h^{-1} ; B. steady state growth at 0.05 h^{-1}

| | Carbon | | Electron | | ATP | |
|--------------|---------------------------|------|--|-----|---|------|
| | Cmol/Cmol _X /h | | mol _e /Cmol _X /h | | mol _{ATP_{eq}} /Cmol _X /h | |
| | A | B | A | B | A | B |
| Source | 0.58 | 0.58 | 2.3 | 2.3 | 0.22 | 0.22 |
| Sinks | 0.49 | 0.40 | 2.0 | 1.6 | 0.22 | 0.08 |
| Recovery (%) | 85 | 69 | 85 | 69 | 99 | 35 |

In phase III (150 – 300 seconds), the discrepancy between the carbon/electron sources and sinks continued to increase. At 300 seconds, 29% of the consumed glucose was channeled into unknown sinks. Fermentative metabolism became the predominant carbon and electron sink, with ethanol, acetate and the related CO₂ evolution comprising 36% of the total glucose uptake at the end of phase III. The ATP balance shows a large energy surplus, which indicates that the unknown carbon/electron sink might be related to yet unquantified energy consuming processes.

One of such processes is growth, which was assumed constant during the transient for the balance calculations. An increase in anabolic fluxes might therefore represent a potential sink for the unaccounted carbon, electrons and the surplus ATP equivalents. However, such an increase would produce a practically undetectable change in the biomass dry weight: assuming that all unaccounted carbon is converted into biomass components, the dry weight would merely increase by about 0.20 g_{DW}/l, compared to a steady state biomass concentration of 15 g_{DW}/l.

To quantify possible changes in anabolic fluxes in phase III, we employed a macroscopic reaction network model assuming a quasi steady state in phase III (see Appendix for details). This type of generic models has been successfully applied to describe catabolic and anabolic processes in transients over several hours (Duboc et al., 1998). The model stoichiometry is given in Table A1 in the Appendix. The anabolic flux was calculated from estimated extracellular fluxes at the end of phase III (Table A1 in the Appendix) and corresponds to a growth rate of 0.14 h^{-1} , compared to a steady state growth rate of 0.05 h^{-1} . Thus, increased anabolic fluxes is consistent with the measured extracellular fluxes and the known macroscopic stoichiometry.

The measured glucose uptake rate and the ATP balance have not been utilized in the above flux calculation, hence, these can be used as an independent check. Two different growth rates were evaluated, i.e. accelerated growth at 0.14 h^{-1} or steady state growth at 0.05 h^{-1} . Table 7.7 shows that assuming increased growth rate, both the carbon and electron recovery improved by 16% to 85% at the end of phase III. In addition, increased growth rate can be energetically well sustained, since the ATP production now closely matches its consumption, while an unchanged growth rate would lead to a large ATP surplus. The latter contradicts with merely a slight increase in the measured ATP level.

It can be calculated from the carbon and electron balances that the remaining

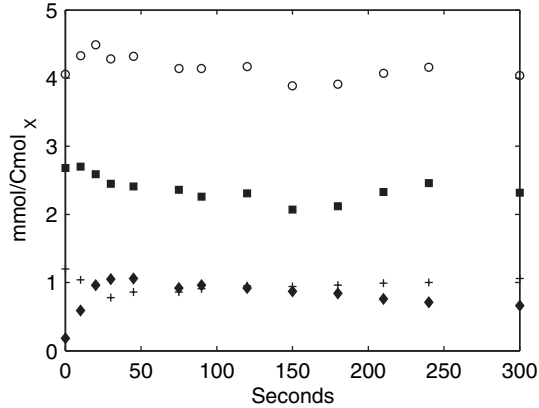


Figure 7.9 Intracellular phosphate balance. Total phosphate present in measured intracellular metabolites: o; Pi: ■; phosphate present in adenine nucleotides: +; phosphate present in metabolites other than adenine nucleotides: ◆.

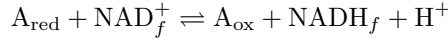
unknown carbon sinks (about 15% of the total glucose uptake) have an average degree of reduction of 4.17 per Cmol. Taking together with the small gap in the ATP balance, the remaining carbon sinks might involve less energy consuming pathways that are neither highly oxidative nor highly reductive. A candidate carbon sink could be glycogen (degree of reduction per Cmol = 4.0) which has a rather low energy requirement ($0.33 \text{ mol}_{\text{ATP}}/\text{Cmol}_{\text{glycogen}}$, compared to $1.67 \text{ mol}_{\text{ATP}}/\text{Cmol}_X$ for biomass). In view of the high glycogen content present in yeast cells, i.e. $650 \mu\text{mol}/\text{g}_{\text{DW}}$ (Guillou et al., 2004), complete conversion of the remaining unknown carbon sinks to glycogen would increase its cellular content by a mere 7%. This solicits highly accurate methods for glycogen quantification.

In addition to the carbon, electron and ATP balances, a balance for intracellular phosphate was constructed using measured concentrations of adenine nucleotides, phosphorylated metabolites and intracellular Pi. Fig. 7.9 shows that increased concentrations of phosphorylated metabolites are compensated by decreased concentrations of adenine nucleotides and Pi, leading to a relatively constant total intracellular concentration of phosphate during the transient. This further implies the absence of possible hydrolysis of intracellular polyphosphate during the transient for additional energy generation.

***In vivo* equilibria and NADH/NAD⁺ ratio** Inspection of the available metabolite measurements provides direct insight into a number of *in vivo* equilibria during the transient. The metabolite pairs G6P/F6P and Fum/Mal have highly correlated concentration changes, indicating that the reversible reactions catalyzed by phosphoglucose isomerase and fumarase are close to equilibrium. In contrast, the reversible reaction catalyzed by enolase was further displaced from equilibrium after the glucose pulse, due to a more pronounced decrease of the concentration of PEP than 2PG+3PG (see Fig. 7.4).

In a number of reversible reactions, free NADH and NAD⁺ participate as cofac-

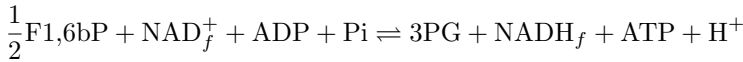
tors, e.g.



If these reactions operate close to equilibrium *in vivo* and the intracellular pH does not significantly change, the ratio between free NADH and free NAD⁺ can be indirectly estimated from the equilibrium constants and measured reactant concentrations, according to the so-called metabolite indicator method (Williamson et al., 1967):

$$\frac{[\text{NADH}_f]}{[\text{NAD}_f^+]} = K_{eq} \frac{[A_{\text{red}}]}{[A_{\text{ox}}]}$$

where K_{eq} stands for the equilibrium constant. The cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$ can be estimated using the lumped reaction catalyzed by the cytosolic enzymes FBA, TPI, GAPDH, PGK and PGM, which convert F1,6bP to 3PG:



The ΔG 's of the individual steps of this lumped reaction are close to 0 under physiological conditions (Voet and Voet, 2004). Assuming equilibrium of this lumped reaction, it can be shown that the following combination of metabolite concentrations is constant:

$$\begin{aligned} & \frac{[2\text{PG} + 3\text{PG}]}{[\text{F1,6bP}]^{1/2}} \frac{[\text{ATP}]}{[\text{ADP}]} \frac{[\text{NADH}_f]}{[\text{NAD}_f^+]} \frac{[\text{H}^+]}{[\text{Pi}]} \\ &= K_{eq,\text{FBA}}^{1/2} K_{eq,\text{TPI}}^{1/2} K_{eq,\text{GAPDH}} K_{eq,\text{PGK}} (1 + K_{eq,\text{PGM}}) \end{aligned}$$

While F1,6bP, 2PG and 3PG can be considered mainly cytosolic, the measured concentration of ATP, ADP and Pi represent cell-averages. Hence, changes of the *apparent* cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$ relative to the steady state were calculated. Fig. 7.10 shows a strong increase in the apparent cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$

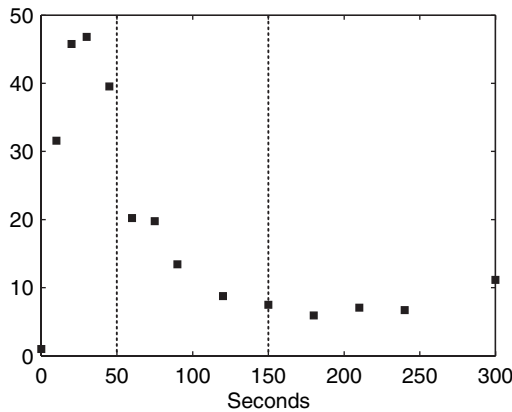
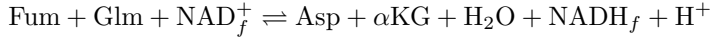


Figure 7.10 Changes of the apparent cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$ relative to the reference steady state, calculated assuming equilibrium of the enzymes FBA, TPI, GAPDH, PGK and PGM.

in phase I, followed by a sharp drop. This trend is comparable to the cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$ estimated by Theobald et al. (1997) in a similar glucose pulse experiment, in which equilibrium of GAPDH and PGK was used with measured cytosolic concentrations of ATP, ADP and Pi.

The NADH/NAD^+ redox couple is also involved in the lumped reaction interconverting C4-metabolites (Fum, Mal, OAA and Asp) by malate dehydrogenase and aspartate aminotransferase:



However, estimation of the $[\text{NADH}_f]/[\text{NAD}_f^+]$ using this lumped reaction is complicated by compartmentation of the enzymes (McAlister-Henn and Thompson, 1987; Minard and McAlister-Henn, 1991; Morin et al., 1992; Verleur et al., 1997) and the metabolites. Interestingly, the total size of the C4-pool (Fum, Mal and Asp) was rather constant (Fig. 7.11), while concentrations of individual pool components changed significantly during the transient (Fig. 7.5). Thus, the in- and out-fluxes of the C4-pool should be well balanced, suggesting that the redistribution of the pool components towards a more reduced status (i.e. increase in the Fum and Mal concentration) might be triggered by increased $[\text{NADH}_f]/[\text{NAD}_f^+]$ in the cytosol and/or mitochondria. This agrees qualitatively with the calculated increase in the cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$ based on equilibrium of lower glycolysis.

***In vivo* flux distribution and kinetic mechanisms in different metabolic regimes**

In phase I, the glucose influx increased rapidly from 0.013 to 0.22 mol/Cmol_X/h at 10 seconds after the pulse. This agrees well with a calculated maximal flux of 0.21 mol/Cmol_X/h based on the Michaelis-Menten kinetic parameters estimated from *zero-trans* experiments (Diderich et al., 1999). This initial increase in q_{Glc} led directly to the accumulation of phosphorylated metabolites and storage material, as well as a strong increase in the cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$. The latter coincides with a sudden increase of q_{O_2} . Immediate secretion of acetate was observed, which is more pronounced than ethanol. This is explained by a much lower affinity

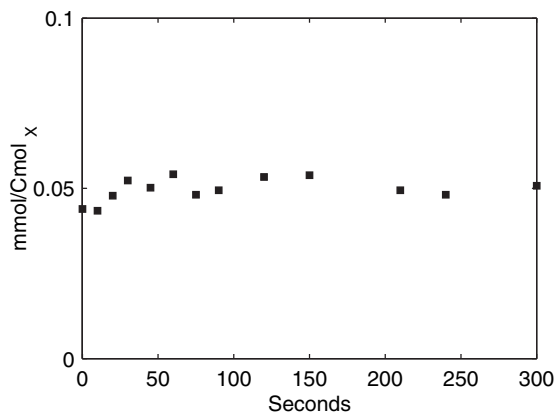


Figure 7.11 Total pool size of C4-metabolites (Fum, Mal and Asp) during the transient.

of alcohol dehydrogenase for acetaldehyde (AA) compared to AADH (Postma et al., 1989), both competing for AA. Furthermore, AADH might be saturated with AA, leading to a rather constant q_{Act} observed during the transient.

In phase I, q_{Glc} sharply dropped directly after the initial increase, which indicates a significant inhibition of the glucose transport by e.g. increased concentration of intracellular glucose (Teusink et al., 1998) and/or G6P (Azam and Kotyk, 1969; Rizzi et al., 1996). The intermediates in upper glycolysis continued to accumulate after the drop of q_{Glc} , which implies that *in vivo* fluxes through the enzymes removing these intermediates, i.e. HK, PFK and the enzymes in lower glycolysis, must have decreased as well, possibly due to complex allosteric regulations. It is well known that ATP is a key effector of HK and PFK (Larsson et al., 2000; Teusink et al., 2000): both enzymes are inhibited at (high) physiological concentrations of ATP. However, a drastically lowered ATP level as observed in phase I can in turn limit the fluxes through the two enzymes (Theobald et al., 1997). Further reduction of HK and PFK activities is presumably due to the elevated level of T6P (see Fig. 7.4), a known inhibitor of HK (Blazquez et al., 1993; Hohmann et al., 1996), and decreased concentration of AMP (after about 20 seconds, see Fig. 7.6), which is a potent activator of PFK (Nissler et al., 1983).

At the pyruvate branch point, flux through the lower glycolysis is partitioned between fermentative and nonfermentative (respiratory) metabolism. This distribution can be analyzed in more detail utilizing the estimated *in vivo* q_{O_2} and q_{CO_2} . The q_{O_2} and q_{CO_2} associated with nonfermentative metabolism are given by:

$$q_{\text{CO}_2, \text{nonferm}} = q_{\text{CO}_2} - (q_{\text{EtOH}} + q_{\text{Act}}), \quad q_{\text{O}_2, \text{nonferm}} = q_{\text{O}_2} - \frac{1}{2}q_{\text{Act}} + \frac{1}{2}q_{\text{Glyc}}$$

assuming a NADP⁺-dependent AADH. In phase I, both the $q_{\text{O}_2, \text{nonferm}}$ and $q_{\text{CO}_2, \text{nonferm}}$ decreased after an initial increase (Fig. 7.12). As $q_{\text{O}_2, \text{nonferm}}$ corresponds to the oxidation of NADH generated in lower glycolysis and the TCA cycle, the observed decrease indicates a drop of the flux through both pathways. A drop of the lower glycolytic flux agrees with the estimated decrease in the cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$ after approximately 30 seconds. A drop of the TCA cycle flux is most likely reflected by the decrease in $q_{\text{CO}_2, \text{nonferm}}$, as the TCA cycle is a major source of $q_{\text{CO}_2, \text{nonferm}}$, supplying 82% of the total q_{CO_2} in the steady state. Changes in pyruvate concentration closely corresponded to the estimated $q_{\text{CO}_2, \text{nonferm}}$, which led further to similar changes observed in most TCA cycle intermediates (Fig. 7.5). As the *in vitro* K_m for pyruvate oxidation by intact mitochondria (0.3 mM, van Urk et al. 1989) is much higher than the pyruvate concentrations encountered in this study, the observation implies that the TCA cycle flux is able to respond quickly to concentration changes of its precursor. Alternatively, a decrease of the TCA cycle flux might be partially due to inhibition of the TCA cycle enzymes (e.g. the pyruvate dehydrogenase complex, see e.g. Pronk et al. 1996) by elevated mitochondrial $[\text{NADH}_f]/[\text{NAD}_f^+]$, as a result of shuttle mechanisms that transport excess cytosolic NADH across the inner mitochondrial membrane (Bakker et al., 2001).

In phase I, it was observed that a considerable amount of carbon is directed to storage carbon metabolism. This is in contrast to previously reported degradation of trehalose and glycogen for derepressed and chemostat grown *S. cerevisiae* upon glucose excess, observed over a time span of hours (van der Plaats, 1974; van Urk et al.,

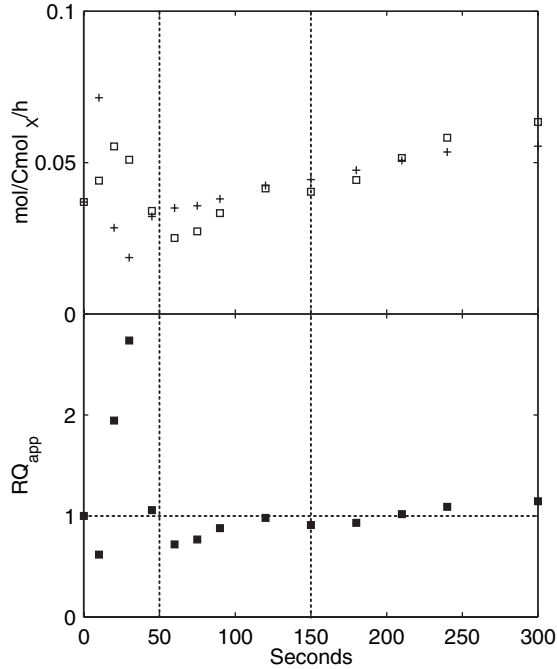


Figure 7.12 Estimated nonfermentative q_{O_2} (+) and q_{CO_2} (□) (upper panel) and the apparent RQ (lower panel).

1990). The disagreement could be attributed to a mutation in adenylate cyclase in the particular CEN.PK-derived strain used in this study, which largely abolishes the glucose-induced increase in cAMP concentration (Vanhalewyn et al., 1999). A rather constant cAMP concentration was indeed measured in our experiment (Kresnowati, unpublished results), in contrast to a strong increase observed with a different *S. cerevisiae* strain upon glucose excess (Vaseghi et al., 2001). As neutral trehalase (the major trehalose-degrading enzyme) is activated through cAMP-dependent protein kinase A (Uno et al., 1983), *S. cerevisiae* strains carrying this mutation are indeed known to exhibit glucose-induced accumulation of trehalose and, to a lesser extent, glycogen (Vanhalewyn et al., 1999).

In phase I, an initial decrease of the 6PG concentration was observed. This is in contrast to an increase found in a similar glucose pulse experiment (Vaseghi et al., 1999). The disagreement might be attributed to differences in the *S. cerevisiae* strain, growth rate and carbon source applied in the two studies. Notably, glucose is the sole carbon source in Vaseghi et al. (1999), while a small amount of ethanol was supplied in this study, which might yield additional NADPH via the NADP⁺-AADH and hence, might result in a lower steady state PP split ratio (see Material and Methods). The drop of the 6PG concentration in our experiment is possibly due to inhibition of G6PDH by excess NADPH from NADP⁺-AADH that accompanied the sudden increase of q_{Act} (Fig. 7.3).

In phase II, the concentrations of the upper glycolytic intermediates continued to decrease even after the increase of q_{Glc} , which implies that the *in vivo* fluxes

though HK and PFK must have increased despite the decreased concentrations of their substrates. This is presumably due to stimulatory effects of recovered ATP and the elevated F2,6bP level (Fig. 7.7), the latter being a strong activator of PFK (Avigad, 1981). The accumulation of F2,6bP observed in this study seems not to have been induced by cAMP, which implicates PKA-independent regulation of phosphofructokinase 2 activity, by e.g. other protein kinases (Dihazi et al., 2003) or allosteric mechanisms (Hofmann et al., 1989).

In phase II, a pronounced increase in q_{EtOH} followed the increase of the pyruvate concentration and leads, together with an increased q_{O_2} , to a lowered cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$. During phase II and III, the $q_{\text{O}_2, \text{nonferm}}$ and $q_{\text{CO}_2, \text{nonferm}}$ continued to increase, up to 1.5 respectively 1.7 fold of the steady state values, while their ratio, given by the apparent RQ, remained close to 1 (Fig. 7.12). This indicates an increased flux towards the TCA cycle and corresponds to increased biosynthesis as calculated by the macroscopic reaction network model. Onset of anabolic processes does not seem to be driven by increased concentrations of amino acids as direct precursors, as they were not significantly affected by the glucose pulse. In addition, during a prolonged chemostat cultivation of *S. cerevisiae*, the growth rate does not change while concentrations of amino acids decrease manifold (Mashego et al., 2005). However, the possibility remains that, due to compartmentation of amino acids, changes in the cytosolic or mitochondrial fractions are masked by the presence of a large vacuolar amino acid pool, metabolically separated from other amino acid pools (Kitamoto et al., 1988; Wiemken and Durr, 1974). Another possible signal for accelerated growth might be an improved cellular energetic status. Although the energy charge is only slightly higher in phase III compared to the steady state, the individual ratios of ATP/ADP and ATP/AMP increase to nearly 2 times their steady state value (not shown). Increased growth further implies an overcapacity of ribosomal proteins at the applied growth rate of 0.05 h^{-1} , as the amount of ribosomes for protein synthesis is likely to remain constant during the relatively short transient of 300 seconds. The overcapacity indicates a low ribosomal efficiency, which has been observed at low specific growth rates both in prokaryotes (Ingraham et al., 1983) and eukaryotes (Bushell and Bull, 1999; Sturani et al., 1979).

In conclusion, by using carbon, electron and ATP balances, we have quantified and analyzed in detail the metabolic responses of *S. cerevisiae* to a glucose pulse, from which three distinct metabolic regimes were recognized during the 300-seconds transient. In phase I (0–50 seconds after the pulse), a high initial glucose influx leads to accumulation of glycolytic and storage metabolites, as well as a large increase in the cytosolic $[\text{NADH}_f]/[\text{NAD}_f^+]$ and q_{O_2} . The accumulation of phosphorylated compounds and storage material causes an energy shortage, which leads to a temporary decrease of the glycolytic and TCA flux and q_{O_2} . It is remarkable that in phase I no significant amount of ethanol is produced. In phase II (50–150 seconds), the fermentative metabolism sets in. The q_{Glc} and q_{EtOH} increased 3 and 8 fold compared to the end of phase I, respectively, which is likely related to the stimulatory effect of increased F2,6bP concentration. A valuable information source is the calculated $q_{\text{O}_2, \text{nonferm}}$ and $q_{\text{CO}_2, \text{nonferm}}$, which point to a progressive increase in the TCA cycle flux. Finally, in phase III (150–300 seconds) a quasi steady state is established, in which about 30% the glucose influx is likely to be channeled into anabolism, which corresponds to an increased growth rate of 0.14 h^{-1} .

References

- Avigad G. 1981. Stimulation of yeast phosphofructokinase activity by fructose 2,6-bisphosphate. *Biochem Biophys Res Commun* 102: 985–91.
- Azam F, Kotyk A. 1969. Glucose-6-phosphate as regulator of monosaccharide transport in baker's yeast. *FEBS Lett* 2: 333–5.
- Bakker BM, Overkamp KM, van Maris AJ, Kotter P, Luttik MA, van Dijken JP, Pronk JT. 2001. Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* 25: 15–37.
- Blazquez MA, Lagunas R, Gancedo C, Gancedo JM. 1993. Trehalose-6-phosphate, a new regulator of yeast glycolysis that inhibits hexokinases. *FEBS Lett* 329: 51–4.
- Bloemen HJJ, Wu L, van Gulik WM, Heijnen JJ, Verhaegen MHG. 2003. Reconstruction of the O₂ uptake rate and CO₂ evolution rate on a time scale of seconds. *AIChE J* 49: 1895–908.
- Bushell ME, Bull AT. 1999. Sporulation at minimum specific growth rate in *Aspergillus nidulans* chemostat culture predicted using protein synthesis efficiency estimations. *J Basic Microbiol* 39: 293–8.
- Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M. 2002. Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol Bioeng* 79: 53–73.
- Claire RL. 2000. Positive ion electrospray ionization tandem mass spectrometry coupled to ion-pairing high-performance liquid chromatography with a phosphate buffer for the quantitative analysis of intracellular nucleotides. *Rapid Commun Mass Spectrom* 14: 1625–34.
- Diderich JA, Schepper M, van Hoek P, Luttik MA, van Dijken JP, Pronk JT, Klaassen P, Boelens HF, de Mattos MJ, van Dam K, Kruckeberg AL. 1999. Glucose uptake kinetics and transcription of HXT genes in chemostat cultures of *Saccharomyces cerevisiae*. *J Biol Chem* 274: 15350–9.
- Dihazi H, Kessler R, Eschrich K. 2003. Glucose-induced stimulation of the ras-cAMP pathway in yeast leads to multiple phosphorylations and activation of 6-phosphofructo-2-kinase. *Biochemistry* 42: 6275–82.
- Duboc P, von Stockar U, Villadsen J. 1998. Simple generic model for dynamic experiments with *Saccharomyces cerevisiae* in continuous culture: decoupling between anabolism and catabolism. *Biotechnol Bioeng* 60: 180–9.
- Guillou V, Plourde-Owobi L, Parrou JL, Goma G, Francois J. 2004. Role of reserve carbohydrates in the growth dynamics of *Saccharomyces cerevisiae*. *FEMS Yeast Res* 4: 773–87.
- Herwig C, Marison I, von Stockar U. 2001. On-line stoichiometry and identification of metabolic state under dynamic process conditions. *Biotechnol Bioeng* 75: 345–54.
- Hofmann E, Bedri A, Kessler R, Kretschmer M, Schellenberger W. 1989. 6-Phosphofructo-2-kinase and fructose-2,6-bisphosphatase from *Saccharomyces cerevisiae*. *Adv Enzyme Regul* 28: 283–306.
- Hohmann S, Bell W, Neves MJ, Valckx D, Thevelein JM. 1996. Evidence for trehalose-6-phosphate-dependent and -independent mechanisms in the control of sugar influx into yeast glycolysis. *Mol Microbiol* 20: 981–91.

- Ingraham JL, Maaloe O, Neidhardt FC. 1983. Growth of the bacterial cell. Sunderland: Sinauer.
- Kitamoto K, Yoshizawa K, Ohsumi Y, Anraku Y. 1988. Dynamic aspects of vacuolar and cytosolic amino acid pools of *Saccharomyces cerevisiae*. J Bacteriol 170: 2683–6.
- Lange HC, Eman M, van Zuijlen G, Visser D, van Dam JC, Frank J, de Mattos MJ, Heijnen JJ. 2001. Improved rapid sampling for *in vivo* kinetics of intracellular metabolites in *Saccharomyces cerevisiae*. Biotechnol Bioeng 75: 406–15.
- Lange HC, Heijnen JJ. 2001. Statistical reconciliation of the elemental and molecular biomass composition of *Saccharomyces cerevisiae*. Biotechnol Bioeng 75: 334–44.
- Larsson C, Pahlman IL, Gustafsson L. 2000. The importance of ATP as a regulator of glycolytic flux in *Saccharomyces cerevisiae*. Yeast 16: 797–809.
- Mashego MR, Jansen ML, Vinke JL, van Gulik WM, Heijnen JJ. 2005. Changes in the metabolome of *Saccharomyces cerevisiae* associated with evolution in aerobic glucose-limited chemostats. FEMS Yeast Res 5: 419–30.
- Mashego MR, van Gulik WM, Vinke JL, Heijnen JJ. 2003. Critical evaluation of sampling techniques for residual glucose determination in carbon-limited chemostat culture of *Saccharomyces cerevisiae*. Biotechnol Bioeng 83: 395–9.
- Mashego MR, Wu L, Van Dam JC, Ras C, Vinke JL, Van Winden WA, Van Gulik WM, Heijnen JJ. 2004. MIRACLE: mass isotopomer ratio analysis of U-¹³C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. Biotechnol Bioeng 85: 620–8.
- McAlister-Henn L, Thompson LM. 1987. Isolation and expression of the gene encoding yeast mitochondrial malate dehydrogenase. J Bacteriol 169: 5157–66.
- Minard KI, McAlister-Henn L. 1991. Isolation, nucleotide sequence analysis, and disruption of the MDH2 gene from *Saccharomyces cerevisiae*: evidence for three isozymes of yeast malate dehydrogenase. Mol Cell Biol 11: 370–80.
- Morin PJ, Subramanian GS, Gilmore TD. 1992. AAT1, a gene encoding a mitochondrial aspartate aminotransferase in *Saccharomyces cerevisiae*. Biochim Biophys Acta 1171: 211–4.
- Nielsen J, Villadsen J. 1994. Bioreaction Engineering Principles. Kluwer Academic Publishers.
- Nissler K, Otto A, Schellenberger W, Hofmann E. 1983. Similarity of activation of yeast phosphofructokinase by AMP and fructose-2,6-bisphosphate. Biochem Biophys Res Commun 111: 294–300.
- Ostergaard S, Olsson L, Nielsen J. 2001. *In vivo* dynamics of galactose metabolism in *Saccharomyces cerevisiae*: metabolic fluxes and metabolite levels. Biotechnol Bioeng 73: 412–25.
- Parrou JL, Francois J. 1997. A simplified procedure for a rapid and reliable assay of both glycogen and trehalose in whole yeast cells. Anal Biochem 248: 186–8.
- Postma E, Verduyn C, Scheffers WA, Van Dijken JP. 1989. Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*. Appl Environ Microbiol 55: 468–77.
- Pronk JT, Yde Steensma H, Van Dijken JP. 1996. Pyruvate metabolism in *Saccharomyces cerevisiae*. Yeast 12: 1607–33.
- Reuss M, Bajpai RK. 1991. Stirred tank modelling. In: Reed G (ed.), Vol. 4, Measuring

- and Control, Biotechnology, volume 4. VCH Verlag, Weinheim, Germany, 2nd edition, 299–348.
- Rizzi M, Baltés M, Theobald U, Reuss M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model. *Biotechnol Bioeng* 55: 592–608.
- Rizzi M, Theobald U, Querfurth E, Rohrhirsch T, Baltés M, Reuss M. 1996. *In vivo* investigations of glucose transport in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 49: 316–27.
- Schmitz M, Hirsch E, Bongaerts J, Takors R. 2002. Pulse experiments as a prerequisite for the quantification of *in vivo* enzyme kinetics in aromatic amino acid pathway of *Escherichia coli*. *Biotechnol Prog* 18: 935–41.
- Stephanopoulos G, Aristidou AA, Nielsen J. 1998. *Metabolic engineering: principles and methodologies*. Academic Press, San Diego.
- Sturani E, Costantini MG, Martegani E, Alberghina L. 1979. Level and turnover of polyadenylate-containing ribonucleic acid in *Neurospora crassa* in different steady states of growth. *Eur J Biochem* 99: 1–7.
- Teusink B, Diderich JA, Westerhoff HV, van Dam K, Walsh MC. 1998. Intracellular glucose concentration in derepressed yeast cells consuming glucose is high enough to reduce the glucose transport rate by 50%. *J Bacteriol* 180: 556–62.
- Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, Snoep JL. 2000. Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem* 267: 5313–29.
- Theobald U, Mailinger W, Baltés M, Reuss M, Rizzi M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng* 55: 305–16.
- Uno I, Matsumoto K, Adachi K, Ishikawa T. 1983. Genetic and biochemical evidence that trehalase is a substrate of cAMP-dependent protein kinase in yeast. *J Biol Chem* 258: 10867–72.
- van Dam JC, Eman M, Frank J, Lange HC, van Dedem GW, Heijnen JJ. 2002. Analysis of glycolytic intermediates in *Saccharomyces cerevisiae* using anion exchange chromatography and electrospray ionization with tandem mass spectrometric detection. *Anal Chim Acta* 460: 209–18.
- van der Plaats JB. 1974. Cyclic 3',5'-adenosine monophosphate stimulates trehalose degradation in baker's yeast. *Biochem Biophys Res Commun* 56: 580–7.
- van Urk H, Mak PR, Scheffers WA, van Dijken JP. 1988. Metabolic responses of *Saccharomyces cerevisiae* CBS 8066 and *Candida utilis* CBS 621 upon transition from glucose limitation to glucose excess. *Yeast* 4: 283–91.
- van Urk H, Schipper D, Breedveld GJ, Mak PR, Scheffers WA, Van Dijk P. 1989. Localization and kinetics of pyruvate-metabolizing enzymes in relation to aerobic alcoholic fermentation in *Saccharomyces cerevisiae* CBS 8066 and *Candida utilis* CBS 621. *Biochim Biophys Acta* 992: 78–86.
- van Urk H, Voll WSL, Scheffers WA, van Dijken JP. 1990. Transient-state analysis of metabolic fluxes in crabtree-positive and crabtree-negative yeasts. *Appl Environ Microbiol* 56: 281–7.
- Vanhalewyn M, Dumortier F, Debast G, Colombo S, Ma P, Winderickx J, Van Dijk

- P, Thevelein JM. 1999. A mutation in *Saccharomyces cerevisiae* adenylate cyclase, Cyr1K1876M, specifically affects glucose- and acidification-induced cAMP signalling and not the basal cAMP level. *Mol Microbiol* 33: 363–76.
- Vaseghi S, Baumeister A, Rizzi M, Reuss M. 1999. *In vivo* dynamics of the pentose phosphate pathway in *Saccharomyces cerevisiae*. *Metab Eng* 1: 128–40.
- Vaseghi S, Macherhammer F, Zibek S, Reuss M. 2001. Signal transduction dynamics of the protein kinase-A/phosphofruktokinase-2 system in *Saccharomyces cerevisiae*. *Metab Eng* 3: 163–72.
- Verduyn C, Stouthamer AH, Scheffers WA, van Dijken JP. 1991. A theoretical evaluation of growth yields of yeasts. *Antonie Van Leeuwenhoek* 59: 49–63.
- Verleur N, Elgersma Y, Van Roermund CW, Tabak HF, Wanders RJ. 1997. Cytosolic aspartate aminotransferase encoded by the AAT2 gene is targeted to the peroxisomes in oleate-grown *Saccharomyces cerevisiae*. *Eur J Biochem* 247: 972–80.
- Visser D, van Zuylen GA, van Dam JC, Eman MR, Proll A, Ras C, Wu L, van Gulik WM, Heijnen JJ. 2004. Analysis of *in vivo* kinetics of glycolysis in aerobic *Saccharomyces cerevisiae* by application of glucose and ethanol pulses. *Biotechnol Bioeng* 88: 157–67.
- Voet D, Voet JG. 2004. *Biochemistry*. Wiley, Hoboken, 3rd edition.
- Vriezen N, Romein B, Luyben KCAM, van Dijken JP. 1997. Effects of glutamine supply on growth and metabolism of mammalian cells in chemostat culture. *Biotechnol Bioeng* 54: 271–86.
- Wiechert W. 2002. Modeling and simulation: tools for metabolic engineering. *J Biotechnol* 94: 37–63.
- Wiemken A, Durr M. 1974. Characterization of amino acid pools in the vacuolar compartment of *Saccharomyces cerevisiae*. *Arch Microbiol* 101: 45–57.
- Williamson DH, Lund P, Krebs HA. 1967. The redox state of free nicotinamide-adenine dinucleotide in the cytoplasm and mitochondria of rat liver. *Biochem J* 103: 514–27.
- Wu L, Lange HC, Van Gulik WM, Heijnen JJ. 2003. Determination of *in vivo* oxygen uptake and carbon dioxide evolution rates from off-gas measurements under highly dynamic conditions. *Biotechnol Bioeng* 81: 448–58.

Appendix

The stoichiometry of the macroscopic model is given in Table A1, which contains 6 processes, of which 4 fluxes (q_{EtOH} , q_{Act} , q_{glycerol} and q_{O_2}) were estimated at the end of phase III from metabolite concentration profiles (see also Fig. 7.3). These are given in Table A1 together with the estimated overall fluxes q_{Glc} and q_{CO_2} . The remaining 2 unknown fluxes $q_{\text{anabolism}}$ and $q_{\text{catabolism}}$ can thus be calculated using 2 metabolite balances, for which CO_2 and NADH were chosen assuming a constant intracellular NADH concentration.

Table A1 Stoichiometry of the macroscopic model and estimated extracellular fluxes at the end of phase III

| Flux | mol/Cmol _X /h | Macroscopic process |
|--------------------------|--------------------------|---|
| q_{EtOH} | 0.081 | $\frac{1}{2}\text{Glc} \rightarrow \text{EtOH} + \text{CO}_2 + \text{ATP}$ |
| q_{Act} | 0.008 | $\text{Glc} \rightarrow \text{Act} + \text{CO}_2 + \text{ATP} + \text{NADH} + \text{NADPH}^b$ |
| q_{Glyc} | 0.009 | $\text{Glc} + \text{ATP} + \text{NADH} \rightarrow \text{Glyc}$ |
| $q_{\text{anabolism}}^a$ | | $0.2\text{Glc} + 1.65\text{ATP} \rightarrow \text{X} + 0.24\text{CO}_2 + 0.38\text{NADH}$ |
| $q_{\text{catabolism}}$ | | $\text{Glc} \rightarrow 6\text{CO}_2 + 12\text{NADH} + 2\text{ATP}$ |
| q_{O_2} | -0.061 | $\text{O}_2 + 2\text{NADH} \rightarrow 2\text{ATP}^c$ |
| q_{CO_2} | 0.157 | |
| q_{Glc} | -0.096 | |

^a The stoichiometry of the anabolic reaction is adapted from Verduyn et al. (1991).

^b For simplicity NADPH was not balanced.

^c An effective P/O ratio of 1 is assumed, according to Verduyn et al. (1991).

Chapter 8

Future Directions

The ultimate aim of kinetic analysis of organisms is to obtain a full, validated kinetic model of the metabolic (sub-) system of interest, which can be used in the metabolic engineering cycle. Although enzyme kinetics has been studied for decades, research on *in vivo* kinetics in microorganisms is still in its infancy. Major accomplishments in the past ten years include the establishment of rapid sampling methods, greatly increased access to intracellular metabolite concentrations, as well as the development of modeling tools, such as approximative kinetic functions and optimization methods, as discussed in the introduction of this thesis. Despite the intensive research efforts, no validated *in vivo* kinetic model has yet been obtained for the model microorganisms under study (i.e. *S. cerevisiae* and *E. coli*), which illustrates the great amount of work needed to obtain quantitative understanding of extreme complex metabolic systems. The work described in this thesis represents advances towards a quantitative kinetic model of the primary metabolism in *S. cerevisiae*. However, a number of issues, both in experimental and modeling aspects, emerge from the described research and will need to be addressed in future research. These are briefly discussed below.

Compartmentation An essential problem for the kinetic modeling of eukaryotic microorganisms, such as *S. cerevisiae*, is compartmentation. Cell-averaged metabolite concentrations are measured with the experimental procedures described in this thesis. Available selective disruption and extraction techniques, such as the digitonin method, cavitation and lyophilization (Soboll et al., 1979; Zuurendonk, 1979) show poor agreement among each other and are hence unreliable. Theobald et al. (1994) reported a method for measuring adenine nucleotides in the yeast cytosol and mitochondria, respectively. The applicability of this method to a wider range of metabolites in microbial cells has however not been demonstrated. In summary, the direct measurement of intracellular metabolite concentrations in the compartments is challenging.

Alternatively, cell-averaged concentrations might be used to estimate kinetic parameters in a compartmentalized model, with relevant transport steps between the compartments. Considering the fact that there are 35 (putative) mitochondrial transport proteins in yeast and that many of them transport a range of metabolites

(Palmieri et al., 2000), modeling of transport processes will require a large number of kinetic parameters that are likely hard to identify. Multiple pulse experiments, preferably using mutants affected in the transport steps or in mitochondrial function, might be needed for adequate parameter estimation.

Free NAD^+/NADH and $\text{NADP}^+/\text{NADPH}$ ratio The free NAD^+/NADH and $\text{NADP}^+/\text{NADPH}$ ratios are indicators of the cellular redox status and influence numerous reactions in the metabolic system. Direct determination of these ratios is complicated by a large amount of enzyme-bound NADH respectively NADP^+ (Tischler, 1977), compartmentation of the nicotinamide nucleotides (Reich and Sel'kov, 1981) and the presence of multiple redox shuttles in *S. cerevisiae* (Bakker et al., 2001). The metabolite indicator method is a promising tool to obtain the free NAD^+/NADH and $\text{NADP}^+/\text{NADPH}$ ratios, as shown in chapter 7 of this thesis and by Theobald et al. (1997). A similar indirect approach was employed to estimate cytosolic pH using benzoic acid as a metabolic indicator (Kresnowati, unpublished results). The validity of this approach needs yet to be demonstrated by independent metabolite couples. In addition, suitably chosen indicator reactions are needed to obtain these ratios in the relevant compartments.

Interaction between glycolysis and the storage carbon metabolism It is shown in chapter 7 that yeast cells accumulate trehalose after a glucose pulse; in addition, it has been suggested that formation of carbohydrate, e.g. glycogen, might be a significant carbon sink next to increased growth. Both trehalose and glycogen are present in yeast cells in high concentrations compared to the total amount of consumed glucose after a glucose pulse, leading to difficulties in quantification. Sensitive analytical procedures are required for adequate modeling of the interaction between glycolysis and the storage carbon metabolism. One possibility is to add a pulse of (specifically) ^{13}C -labeled glucose, which will form ^{13}C -labeled storage carbohydrate, thereby providing resolution from the large pools of naturally labeled storage carbohydrate by MS analysis. The exchange fluxes between storage carbohydrate and glycolysis should however be properly taken into account in this type of experiments.

Experimental design, validation and statistical analysis For stimulus response experiments, the choices of external stimuli are often limited, as not all substances can be transported through the cell membrane. Nevertheless, the size of perturbation, the observation time window, the number of samples and the sampling frequency can still be optimized to maximize the relevant information content for parameter estimation. A qualitative guideline drawn from the observations in chapter 7 is that a shorter time frame should be chosen (i.e. less than 100 seconds) with more measurements in the most dynamic phase (i.e. phase I).

To *validate* a kinetic model obtained from e.g. a glucose pulse experiment, different (combinations of) external stimuli can be used. A problem hereby is that different metabolic sub-networks might be perturbed in validation experiments than in the glucose pulse experiment. In addition, mutant strains can be used for validation experiments using the same glucose pulse. A problem hereby is the changed reference enzyme capacities in the mutants. Alternatively, the measurements during a stimulus response experiment can be split into two datasets for parameter estima-

tion and model validation, respectively. The two datasets might contain consequent time trajectories or different sets of measured metabolite concentrations.

Statistical properties of the obtained kinetic parameters can be investigated. Given the estimated properties of measurement errors, the statistical analysis will yield the (co)variance matrix of the estimated kinetic parameters with their confidence intervals. The statistical treatment can be greatly simplified if the linear regression method, proposed in (Kresnowati et al., 2005), is applied to estimate the elasticities. Alternatively, the covariance matrix can be obtained by linearizing the kinetic model around the estimated kinetic parameters. The statistical analysis will identify kinetic parameters that are not sensitive to the measurements or having high correlations with others. Such information can be used to for improved experimental design or model reduction (e.g. elimination of insensitive parameters from the estimation procedure).

Identification of *in vivo* interactions The modulation matrix contains information on enzyme-metabolite interactions, which is a prerequisite for estimating kinetic parameters. Currently, the only information source of these interactions is *in vitro* biochemical studies of enzymes, where the influence of a limited number of metabolites is investigated. It can be expected that 1) more interactions exist *in vivo*, as there are far more metabolites present in cells than tested *in vitro* and 2) the *in vitro* predicted interaction might not prevail *in vivo*. Thus, methods for systematic identification and validation of *in vivo* interactions will be required. In terms of MS analysis, the range of metabolites will need to be expanded; moreover, unknown peaks in LC-MS/MS analysis, which show significant differences between the perturbed and the reference metabolic network, will need to be identified. Families of kinetic models, each having a different modulation structure, can be used to evaluate the significance of hypothesized interactions by comparing the goodness of fit to the experimental data (Haunschild et al., 2005).

Another related issue is to understand interactions in a systems manner, that is, how pathway properties arise from kinetic properties of individual enzymes and their interactions. With the identified *in vivo* interactions one might hope to resolve the *in vivo* - *in vitro* paradigm, i.e. how do interactions translate, when parts of an integrated system are studied in isolate (i.e. *in vitro*), into observed biological functions when they are put back together (i.e. *in vivo*).

Improvement and application of the lin-log kinetics The lin-log kinetics is a useful tool for *in vivo* kinetic analysis, as shown in chapter 4, 5, and 6 in this thesis and by Visser et al. (2004) and Heijnen et al. (2004). A number of possible further developments are envisaged below.

First, the kinetic format can be extended to describe gene expression, protein synthesis and modulation of enzyme capacities, thereby incorporating transcriptome and proteome data, obtained e.g. from a longer transient in a stimulus response experiment. This is a challenging task, due to both the lack of *a priori* information of regulation within and between these various ‘omes’ and the lack of experimental data; for example, levels of phosphorylated proteins are difficult to quantify.

Second, in the current lin-log kinetic format, reaction rates change monotonously with metabolite concentrations. *In vitro*, substrate inhibition is observed, such as for phosphofructokinase, which is inhibited by ATP at high physiological concentrations

(Teusink et al., 2000). Thus, further development of the lin-log kinetics is necessary to address complex allosteric mechanisms of crucial enzymes.

Finally, the robustness of the lin-log kinetics can be tested with dynamic experimental data. Chapter 4 and 6 shows that the lin-log kinetics can be successfully applied to multiple steady state data. Methods for elasticities estimation from time series of metabolite concentrations, such as can be obtained from a stimulus response experiment, has been developed for the lin-log kinetics (Kresnowati et al., 2005). Recently, Mauch et al. (2004) demonstrated the feasibility of large scale parameter estimation using the lin-log kinetics from metabolite time series obtained from a stimulus response experiment on *E. coli*. It will thus be interesting to apply the metabolome measurement presented in Chapter 7, as well as the constraints derived in Chapter 6, for the construction of a lin-log kinetic model of the primary metabolism of *S. cerevisiae*. Once such a model is available, it can be used to test the functional genomics strategy presented in Chapter 5.

References

- Bakker BM, Overkamp KM, van Maris AJ, Kotter P, Luttik MA, van Dijken JP, Pronk JT. 2001. Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* 25: 15–37.
- Haunschild MD, Freisleben B, Takors R, Wiechert W. 2005. Investigating the dynamic behavior of biochemical networks using model families. *Bioinformatics* 21: 1617–25.
- Heijnen JJ, Van Gulik WM, Shimizu H, Stephanopoulos G. 2004. Metabolic flux control analysis of branch points: an improved approach to obtain flux control coefficients from large perturbation data. *Metab Eng* 6: 391–400.
- Kresnowati MT, van Winden WA, Heijnen JJ. 2005. Determination of elasticities, concentration and flux control coefficients from transient metabolite data using linlog kinetics. *Metab Eng* 7: 142–53.
- Mauch K, Schmid JW, Reuss M, Ulmer H. 2004. *In silico* identification of whole cell metabolite dynamics through evolutionary algorithms and parallel computing. In: 5th International conference on systems biology. Heidelberg, Germany.
- Palmieri L, Lasorsa FM, Voza A, Agrimi G, Fiermonte G, Runswick MJ, Walker JE, Palmieri F. 2000. Identification and functions of new transporters in yeast mitochondria. *Biochim Biophys Acta* 1459: 363–9.
- Reich JG, Sel'kov EE. 1981. Energy metabolism of the cell. Academic Press, London.
- Soboll S, Elbers R, Heldt HW. 1979. Metabolite measurements in mitochondria and in the extramitochondrial compartment by fractionation of freeze-stopped liver tissue in nonaqueous media. *Methods Enzymol* 56: 201–6.
- Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, Snoep JL. 2000. Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem* 267: 5313–29.
- Theobald U, Mailinger W, Baltes M, Reuss M, Rizzi M. 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnol Bioeng* 55: 305–16.

- Theobald U, Mailinger W, Rizzi M. 1994. Use of HgCl₂ to investigate dynamic phenomena in yeast cytoplasm. *Biotechnol Techniq* 8: 723–8.
- Tischler ME. 1977. Pyridine nucleotide distributions and enzyme mass action ratios in hepatocytes from fed and starved rats. *Arch Biochem Biophys* 184: 222–36.
- Visser D, Schmid JW, Mauch K, Reuss M, Heijnen JJ. 2004. Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. *Metab Eng* 6: 378–90.
- Zuurendonk PF. 1979. Rapid separation of particulate and soluble fractions from isolated cell preparations (digitonin and cell cavitation procedures). *Methods Enzymol* 56: 207–23.

Summary

Of the thesis: ‘Development and Application of Experimental and Modeling Tools for *In Vivo* Kinetic Analysis in *S. cerevisiae*’ by Liang Wu

Mathematical models of *in vivo* kinetics enhance our understanding of the control of metabolic networks and can be applied in metabolic engineering to direct genetic alterations for desired flux distribution. To obtain *in vivo* kinetic information needed for this type of models, the metabolic system is usually perturbed, which can result in different physiological (quasi) steady states (i.e. multiple steady state experiments), or in a dynamic transient (i.e. stimulus response experiments). The metabolic responses are experimentally determined and subsequently applied for *in vivo* kinetic parameter estimation. In this thesis, a number of novel experimental and modeling methods are developed and applied to study *in vivo* kinetics in the primary metabolism of *S. cerevisiae*.

The O₂ uptake rate (OUR) and CO₂ evolution rate (CER) are two physiologically important responses and are highly dynamic during a stimulus response experiment. So far no suitable tools are available to derive the OUR and CER based on measured O₂ and CO₂ content in the fermentor off-gas, due to delay and distortion effects caused by the fermentor setup. In chapter 2, a black-box model was set up for the fermentor setup to describe these effects. After proper identification of the unknown model parameters, the OUR and CER were obtained with signal processing algorithms. Using this method, distinct dynamic responses of OUR and CER were found after a glucose respectively an ethanol pulse to a chemostat culture of *S. cerevisiae*. The method has been further improved to take into account the variation in off-gas flow rates during a stimulus response experiment (Bloemen et al., 2003).

In addition to the OUR and CER, accurate quantification of the intracellular metabolome is needed for *in vivo* kinetic analysis. The sample processing and analysis, using e.g. liquid chromatography tandem mass spectrometry (LC-MS/MS), involves a number of undesired effects, such as metabolite degradation, ion suppression and operational variations. These effects lead to lowered precision and high method development efforts. In chapter 3, U-¹³C-labeled metabolites were applied as ideal internal standards to correct for the said deleterious effects. These fully-labeled internal standards were prepared with a fed-batch fermentation of *S. cerevisiae* on fully-labeled carbon sources, resulting in a fully-labeled metabolome. Using this method, the analytical precision was considerably improved and method development efforts can be significantly reduced.

A suitable kinetic format is needed to construct mathematical models of the metabolic network from experimental data. The traditional metabolic control analysis (MCA) is based on linearized kinetics. The kinetic parameters are the so-called elasticities, defined as the scaled sensitivity of reaction rates towards metabolite concentrations. However, parameter estimation with MCA proved to be problematic, due to the linear nature of the kinetic format and the inherent nonlinearity in biological systems. A nonlinear, yet structured approximative kinetic format, the linear-logarithmic (lin-log) kinetics, allows adequate description of the metabolic net-

work, as shown in several *in silico* studies. In the lin-log kinetics, the reaction rate is proportional to the enzyme capacity and a linear combination of the logarithms of metabolite concentrations. In chapter 4, the applicability of the lin-log kinetics was examined on experimental data from an *in vitro* multiple steady state experiment. In the presence of large changes in enzyme capacities, metabolite concentrations and fluxes, the lin-log kinetic format led successfully to a consistent set of elasticities and control coefficients.

In chapter 5, the lin-log kinetic format was further applied in a proposed conceptual functional genomic strategy, using solely metabolome data. The function of a gene can be revealed by changing its expression level and analyzing the resulting (flux) phenotype. In the absence of observable flux phenotypes (i.e. ‘silent mutants’), the data-driven FANCY approach has been developed, which utilizes metabolome data for function assignment by comparing the co-response of metabolite concentrations in strains deleted for known and unknown genes (Raamsdonk et al., 2001). In our approach, *in vivo* enzyme activity changes in mutants are predicted from metabolome data and are employed to formulate hypotheses to infer unknown gene functions. Metabolite concentrations, fluxes and *in vivo* enzyme kinetic parameters are needed for our approach, which can all be obtained from metabolome analysis. In an *in silico* case study, it was shown that using the simple lin-log kinetics, enzyme capacity changes in mutants can be accurately predicted, even in ‘silent mutants’.

Dynamic and multiple steady state perturbation experiments and the lin-log kinetics were applied to *in vivo* kinetic analysis of the primary metabolism of *S. cerevisiae*. In chapter 6, *in vivo* kinetics in the primary metabolism of *S. cerevisiae* was investigated through multiple steady states occurring during prolonged chemostat cultivation of *S. cerevisiae* for 70 – 95 generations. The observed large changes in enzyme and metabolite levels, in contrast to the apparently constant fluxes, was explained by two kinetic mechanisms: 1) reversible reactions are near-equilibrium *in vivo* and 2) irreversible reactions are tightly regulated by coordinated changes of metabolic effectors. The measured enzyme and metabolite levels in prolonged chemostats contain however limited information to estimate *in vivo* kinetic parameters; instead only linear constraints between the parameters were obtained. When perturbed by a glucose pulse, the prolonged culture showed a diminished glucose uptake and ethanol excretion, compared to a 10 generation old chemostat culture. This is likely due to a reduced ATP regeneration capacity.

A glucose pulse on a chemostat culture is frequently applied as a stimulus in stimulus response experiments. So far, kinetic models based on glucose pulse experiments often focus on a metabolic sub-network (e.g. glycolysis) and contain simplifying assumptions about anabolic reactions. To understand the changes in both catabolism and anabolism during such experiments, mass, redox and energy balances were constructed in chapter 7, based on comprehensive metabolome measurements in a glucose pulse experiment on *S. cerevisiae*. Three different phases were discerned from the metabolome responses. In phase I (the first 50 seconds), 60% of the glucose taken up by cells accumulated as intracellular phosphorylated compounds and storage metabolites. The resulting energy shortage probably led to a temporary decrease of the glucose influx. In phase II (50 – 150 seconds), fermentative metabolism set in and became the major carbon and electron sink. The glucose influx gradually increased accompanied by a recovery of the energy charge. In phase III (150 – 300 sec-

onds), a quasi steady state in fluxes and metabolite levels was established, in which 29% of the consumed glucose was not recovered by metabolome measurements and might be partially channeled into anabolic processes. In addition, insights in *in vivo* kinetic mechanisms, flux distribution and equilibrium relationships were obtained, which is prerequisite for the construction of kinetic models.

In conclusion, this thesis presents the application of novel experimental and theoretical tools for *in vivo* kinetic analysis, leading to the increased insight into *in vivo* kinetic mechanisms in *S. cerevisiae*. However, a validated *in vivo* kinetic model of the primary metabolism of *S. cerevisiae* has yet not been obtained. A number of present challenges is discussed in chapter 8.

References

- Bloemen HJJ, Wu L, van Gulik WM, Heijnen JJ, Verhaegen MHG. 2003. Reconstruction of the O₂ uptake rate and CO₂ evolution rate on a time scale of seconds. *AIChE J* 49: 1895–908.
- Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG. 2001. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* 19: 45–50.

Samenvatting

Van het proefschrift: 'Development and Application of Experimental and Modeling Tools for *In Vivo* Kinetic Analysis in *S. cerevisiae*' door Liang Wu

Mathematische modellen van *in vivo* kinetiek verbeteren ons begrip van de control van metabolische netwerken en kunnen toegepast worden in metabolic engineering om genetische manipulaties doelen te localiseren, hetgeen resulteert in een verbeterd flux distributie. Om de benodigde *in vivo* kinetische informatie te winnen, perturbeert men vaak de metabolische netwerk. Dit leidt tot verschillende fysiologische (quasi) steady states (d.w.z. meervoudige steady state experimenten), of een dynamische transient (d.w.z. stimulus response experimenten). De metabolische response wordt experimenteel bepaald en vervolgens gebruikt om bijvoorbeeld *in vivo* kinetische parameters te schatten. In deze proefschrift wordt een aantal nieuwe experimentele en modelering methoden ontwikkeld en/of toegepast om de *in vivo* kinetiek van de primaire metabolisme van *Saccharomyces cerevisiae* te bestuderen.

De O₂ opname snelheid (OUR) en CO₂ evolutie snelheid (CER) zijn twee belangrijk fysiologische responses en zijn zeer dynamische tijdens een stimulus response experiment. Tot dusver zijn er geen geschikte methodiek om deze twee grootheden te berekenen uit de gemeten O₂ and CO₂ gehalte in de fermentor off-gas. Dit heeft te maken met de vertraging- en vervormings-effecten van de fermentor setup. In hoofdstuk 2 werden deze effecten beschreven met een black-box model van de fermentor setup. Na de identificatie van model parameters kunnen de OUR en CER berekend worden met signaal verwerkings-algorithm. Met deze methode werden kenmerkende verschillen gevonden tussen de dynamische responses van de OUR en CER, wanneer men een glucose respectievelijk een ethanol pulse gaf aan een koolstof-gelimiteerd chemostaat van *S. cerevisiae*. Deze methode is verder verbeterd om de veranderend volume van het off-gas tijdens zo'n stimulus response experiment in rekening te nemen (Bloemen et al., 2003).

De nauwkeurige quantificatie van intracellulair metabolieten is nodig voor *in vivo* kinetiek analyse. Hiervoor blijkt LC-MS/MS een geschikte meet methode. Een aantal niet gewenste effecten in monster-bereiding en analyse, bijvoorbeeld metaboliet degradatie, ion suppressie en operationeel variaties, kan echter de analyse nauwkeurigheid verlagen en de methode-ontwikkeling vermoeilijken. In hoofdstuk 3 werd U-¹³C-gelabeld metabolieten gebruikt worden als interne standaarden om de genoemde effecten te corrigeren. Deze fully-labeled interne standaarden werden verkregen door *S. cerevisiae* te groeien in een fed-batch aan fully-labeled koolstof bronnen. Dit leidde tot een fully-labeled metabolome. Met deze methode was de analyse precisie aanzienlijk verbeterd en method ontwikkeling significant vereenvoudigd.

Een geschikte kinetiek formaat is nodig om wiskundige modellen van metabolische netwerk te construeren uit experimenteel data. De traditionele metabolic control analysis (MCA) is gebaseerd op gelineariseerd kinetiek. De kinetische parameters zijn de zogenaamde elasticiteiten, gedefinieerd als de geschaald sensitiviteit van reactie snelheden naar metaboliet concentraties. Parameter schatten met MCA kan

echter problemen geven, door de lineaire aard van de onderliggende kinetische formaat en de niet-lineaire natuur van biologische systemen. Een niet-lineaire, doch gestructureerde benaderings- kinetiek formaat, namelijk de lineair-logarithmische (lin-log) kinetiek, is voorgesteld om de metabolische netwerk adequaat te beschrijven, zoals aangetoond is in een aantal *in silico* studies. In de lin-log kinetiek, de reactie snelheden is evenredig met de enzym capaciteit en een lineaire combinatie van logarithmische metaboliet concentraties. In hoofdstuk 4 werd deze kinetiek formaat toegepast op experimenteel data uit een *in vitro* meervoudige steady state experiment, waarin groot veranderingen van fluxen, metaboliet concentraties en enzym capaciteiten voorkomen. Toepassing van de lin-log kinetiek leidde tot een consistent set elasticiteiten en control coefficients.

De lin-log kinetische formaat was verder toegepast in hoofdstuk 5 in een voorgestelde conceptueel functioneel genomics strategie, gebruikmakend van metabolome data alleen. De functie van een gen kan bestudeerd worden door de expressie niveau van deze gen te verandering en de resulterende flux-fenotype te analyseren. In afwezigheid van een observeerbaar flux fenotype ('silent mutants'), de data-driven FANCY aanpak is voorgesteld om met metabolome data de gen-functie te ontrafelen, door het vergelijken van co-response van metaboliet concentraties in knockout mutanten van bekende en onbekende genen (Raamsdonk et al., 2001). In onze aanpak worden veranderingen van *in vivo* enzym capaciteiten berekend uit metabolome data. Deze berekende veranderingen leiden tot zinvolle hypothesen om gen-functies af te leiden. Nodig voor deze aanpak zijn de metaboliet concentraties, fluxen en *in vivo* enzym kinetiek parameters, die allemaal verkregen kunnen worden uit metabolome data. In een *in silico* case studie werd aangetoond dat met de simpele lin-log kinetiek, veranderingen van enzym activiteiten in mutanten kunnen nauwkeurige worden voorgespeeld, zelfs in 'silent mutants'.

Dynamische en meervoudige steady state perturbatie experimenten en de lin-log kinetiek zijn toegepast in de *in vivo* kinetische analyse van de primaire metabolisme van *S. cerevisiae*. In hoofdstuk 6 was *in vivo* kinetiek van de primaire metabolisme in *S. cerevisiae* bestudeerd met meervoudige steady states, die voorkwamen tijdens verlengde chemostaat van *S. cerevisiae* van 70 – 95 generaties, waarin grote veranderingen in enzyme en metabolieten niveau ging gepaard met schijnbare constante fluxen. Dit werd verklaard met twee kinetiek mechanismen: 1) reversibele reacties zijn dichtbij evenwicht *in vivo* en 2) irreversibele reacties zijn sterk gereguleerd door gecoördineerde veranderingen van metaboliet effectoren. De gemeten enzyme en metabolite niveaus in meervoudige steady states bevatten echter gelimiteerd kinetiek informatie om *in vivo* kinetische parameters te schatten, in plaats daarvan werden alleen lineaire constraints verkregen. Na een glucose puls had de verlengde chemostaat een sterk verlaagde glucose opname snelheid en ethanol productie, vergeleken met een 10 generatie oud chemostaat. Dit is waarschijnlijk veroorzaakt door een gereduceerde capaciteit voor ATP regeneratie.

A glucose pulse op een chemostaat cultuur is vaak toegepast als een perturbatie middel in stimulus response experimenten. Kinetische modellen afgeleid uit zulke glucose pulse experimenten zijn meestal gefocuseerd op een metabolische sub-netwerk (bijvoorbeeld glycolyse) en bevatten vereenvoudigende aannames over anabolische gedeelte van de metabolisme. Om veranderingen in de catabolisme en de anabolisme beter te begrijpen werden in hoofdstuk 7 koolstof, electronen en ATP

balansen geconstrueerd met uitgebreide metabolome metingen, nadat een koolstof-gelimiteerde chemostaat van *S. cerevisiae* was verstoord met een glucose puls. Drie verschillende fasen werden gekend van de metabolome responses. In fase I (de eerste 50 seconden) werd 60% van de opgenomen glucose geaccumuleerd in de cellen als gefosphoryleerde verbindingen en opslag materialen. De resulterende energie tekort leide tot een tijdig afname van de glucose influx. In fase II (50 – 150 seconden), fermentatieve metabolisme is aangezet en werd de voornamste koolstof sink. De glucose influx nam weer toe door een verhoogde energie charge. In fase III (150 – 300 seconden), een quasi steady state in fluxen and metaboliet concentraties kwam tot stand, waarin 29% van de opgenomen glucose niet in de metabolome metingen teruggevonden kunnen worden. Deze kookstof zou door anaolische processen geconsumeerd kunnen worden. Daarnaast werd inzicht verkregen in *in vivo* kinetiek mechanismen, flux verdeling en evenwichts-relaties, welke vereist is voor het bouwen van kinetische modellen.

Samenvattend heeft de toepassing van de nieuwe experimentele en modellerings-methode in deze thesis geleid tot een betere inzicht in de primaire metabolisme van *S. cerevisiae*. Een gevalideerd *in vivo* kinetische model is echter nog niet tot stand gekomen. In hoofdstuk 8 werd er over de uitdagingen voor *in vivo* kinetische analyse gediscussieerd.

References

- Bloemen HJJ, Wu L, van Gulik WM, Heijnen JJ, Verhaegen MHG. 2003. Reconstruction of the O₂ uptake rate and CO₂ evolution rate on a time scale of seconds. *AIChE J* 49: 1895–908.
- Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG. 2001. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* 19: 45–50.

List of Publications

Journal articles

- Wu L, Lange HC, van Gulik WM, Heijnen JJ. 2003. Determination of *in vivo* oxygen uptake and carbon dioxide evolution rates from off-gas measurements under highly dynamic conditions. *Biotechnol Bioeng* 81: 448-58.
- Bloemen HJJ, Wu L, van Gulik WM, Heijnen JJ, Verhaegen MHG. 2003. Reconstruction of the O₂ uptake rate and CO₂ evolution rate on a time scale of seconds. *AIChE J* 49: 1895-1908.
- Mashego MR, Wu L, Van Dam JC, Ras C, Vinke JL, Van Winden WA, Van Gulik WM, Heijnen JJ. 2004. MIRACLE: mass isotopomer ratio analysis of U-¹³C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites. *Biotechnol Bioeng* 85: 620-8.
- Wu L, Wang WM, van Gulik WM, Heijnen JJ. 2004. A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. *Eur J Biochem* 271: 3348-59.
- Visser D, van Zuylen GA, van Dam JC, Eman MR, Proll A, Ras C, Wu L, van Gulik WM, Heijnen JJ. 2004. Analysis of *in vivo* kinetics of glycolysis in aerobic *Saccharomyces cerevisiae* by application of glucose and ethanol pulses. *Biotechnol Bioeng* 88: 157-67.
- Wu L, Mashego MR, van Dam JC, Proell A, Vinke JL, Ras C, van Winden WA, van Gulik WM, Heijnen JJ. 2005. Metabolome analysis by isotope dilution mass spectrometry using fully U-¹³C-labeled cell extracts as internal standards. *Anal Biochem* 336: 164-71.
- Wu L, van Winden WA, van Gulik WM, Heijnen JJ. 2005. Application of metabolome data in functional genomics: a conceptual strategy. *Metab Eng* 7: 302-10.
- Wu L, Mashego MR, Proell AM, Vinke JL, Ras C, van Dam JC, van Winden WA, van Gulik WM, Heijnen JJ. *In vivo* kinetics of primary metabolism in *S. cerevisiae* studied through prolonged chemostat cultivation. Accepted for publication in *Metab Eng*.
- Wu L, van Dam JC, Schipper D, Kresnowati MT, Proell AM, Ras C, van Winden WA, van Gulik WM, Heijnen JJ. Analysis of *in vivo* dynamics in *S. cerevisiae* by comprehensive metabolome quantification and mass, electron and ATP balances. Submitted for publication in *Appl Environ Microbiol*.

Poster presentations

- Wu L, Lange HC, van Gulik WM, Heijnen JJ. Dynamic measurements of off-gas signals in perturbed chemostats in a 300 seconds window for *in vivo* kinetics of *S. cerevisiae*. Presented at Metabolic Engineering III-Conference, Colorado Springs, USA, 2000 and at ESBES-3, Copenhagen, Denmark, 2000.
- Wu L, Mashego MR, Vinke JL, van Gulik WM, Heijnen JJ. Prolonged glucose-limited chemostat cultivation affects the metabolic response of *S. cerevisiae* to a glucose pulse. Presented at Metabolic Engineering IV-Conference, Il Ciocco, Italy, 2002.

List of publications

- Wu L, Bloemen HJ, van Gulik WM, Verhaegen MHG, Heijnen JJ. Reconstruction of the dynamic oxygen uptake rate and carbon dioxide evolution rate during a pulse experiment. Presented at Metabolic Engineering IV-Conference, Il Ciocco, Italy, 2002.
- Bloemen HJ, Wu L, van Gulik WM, Verhaegen MHG, Heijnen JJ. A dynamic soft sensor for the O₂ uptake rate and the CO₂ evolution rate on a time scale of seconds. Presented at ESBES-4, Delft, The Netherlands, 2003.
- Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, van Winden WA, van Gulik WM, Heijnen JJ. Metabolome analysis by isotope dilution mass spectrometry using fully-labeled cell extracts as internal standards. Presented at Physiology of Yeasts and Filamentous Fungi 2nd Symposium, Anglet, France, 2004.
- Wu L, van Winden WA, van Gulik WM, Heijnen JJ. Application of metabolome data in functional genomics: an integrative strategy. Presented at Metabolic Engineering-V, Lake Tahoe, USA, 2004.

Oral presentations

- Reconstruction of the dynamic OUR and CER during a pulse experiment. Presented at Advanced Course Technologies for Metabolic Engineering in Biotechnology and Medicine, Carcavelos, Portugal, 2002.
- Experimental Methods for the study of *in vivo* cellular dynamics through environmental perturbation. Presented at DSM scientific meeting, Vaals, The Netherlands, 2003.
- Lin-log kinetic approach: a new theoretical tool to explore metabolic networks. Presented at Kluyver Center Symposium, Garderen, The Netherlands, 2004.
- A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. Presented at Dutch Biotechnology Congress 10, Ede, The Netherlands, 2004.
- Application of metabolome data in functional genomics: an integrative strategy. Presented at ESBES-5, Stuttgart, Germany, 2004.

Curriculum Vitae

Liang Wu was born on July 12, 1975 in Hanzhong, Shaanxi Province, The People's Republic of China. From 1987 to 1993 he attended the Shanghai No.2 Middle School in Shanghai, P. R. China. He then moved to the Netherlands in 1993 for family reunion. From 1993 to 1994 he followed Dutch language courses in James Boswell Institute in Utrecht and at the Delft University of Technology (TU Delft). From 1994 to 1999 he studied chemical engineering at TU Delft. In 1998, he performed his internship at Diosynth BV and worked as a student fellow in the reactor and catalysis engineering group at TU Delft. He carried out his graduation project in the bioprocess technology group, during which he developed methods to reconstruct dynamic oxygen and carbon dioxide conversion rates using measured off-gas concentrations. From 2000 to 2004 he worked as a PhD in the bioprocess technology group under the supervision of Prof. dr. J.J. Heijnen. The research aims to develop experimental and modeling methods for *in vivo* kinetic analysis; the major results are presented in this thesis. Liang Wu is currently employed as a post-doctoral researcher at the Institute of Biotechnology 2, Forschungszentrum Jülich in Germany.

Dankwoord

The conventional placement of the ‘Dankwoord’ section behind the main body of the thesis fails to reflect a basic causal relationship, that is, this thesis wouldn’t be realized without the effort and input of many others in the first place.

My thank first goes to my promoter Prof. Sef Heijnen. There are different kinds of professors, and I really appreciate that Sef is one that truly ‘supervises’ and with great interest, involves in the scientific research. Although almost always busy, he never let a chance go to pop into the PhDs’ offices to have a scientific discussion. In addition to being an insightful scientist, Sef is an inspiring teacher and mentor, from whom I greatly benefited during my years in Delft. Sef, thanks for your trust, support, as well as exchange and supply of many ideas during my research.

I would like to acknowledge Walter van Gulik for supervision and support on the daily working- and lab-front. Many chapters in this thesis would be less readable and less structured, were it not the effort of Walter, and later also Wouter van Winden, who critically read and corrected them.

Many helping hands have contributed to the experimental results presented in this thesis. Hans Lange was the first who guided me into the world of yeast fermentation and dynamic experiments. Ko Vinke was extremely helpful in all fermentation experiments and kept the fermentation lab running smoothly during his presence. Although heavily occupied, Rob Kerste and Dirk Geerts provided adequate technical support; I have to thank you in particular for the dedicated Labview interface for off-gas/DO data logging and system identification experiments. Rapid sampling experiments will never be rapid enough without the rapid sampling team. My thanks to the team members Walter, Wouter, Ko, Mlawule Mashego, Roelco Kleijn, Uly Nasution, PeNia Kresnowati and Angie Proell, who regularly participated in this stressful activity. I am most grateful to Angie and Cor Ras in the LC-MS/MS team, led by Jan van Dam, for providing high quality and reliable measurements. Jan, Angie en Cor: BEDANKT!

The regular project meeting was attended by members of the yeast group, a.o. Prof. Jack Pronk, and members of DSM research, whom I would like to acknowledge for their interest and feedback. I thank Dick Schipper of DSM in particular for providing the trehalose measurement. I also sincerely thank Klaus Mauch in University of Stuttgart for collaboration in the ‘Insilico’ project, although the intermediate results have not been included in this thesis. I am confident that the future is bright for ‘Insilico’. My thanks also go to the two students I have coached, PeNia and Weiming Wang.

Less scientifically related but not less important are the supporting services in the Kluiver lab. I would like to thank in particular Sjaak, Carla, Jenny and Miranda of the secretary, Pascal and Astrid in the ‘keuken’, systeembeheerder Hans and Marcel and Arno of the werkplaats.

So far for people I worked with and now some room for people I ‘lived’ with. First come my family: My parents and families in Shanghai: grandma, auntie and uncle. I can’t possibly tell you how grateful I am to you. Dad and Jet: you have provided a place I could call a second ‘home’ (and not to forget Bingbing en dikke

Dankwoord

Tonny), where I can fall back for ‘gezelligheid’, but also for recovering from stress and tiredness. Thanks! Words would simply fail me . . . Next I would like to thank *the architect* and my girlfriend, Wenqian Shi, for sharing all ups and downs of life with me.

Officemates (Hans, Preben Krabben, Hayco Bloemen, Xiaonan Li, Frederik Aboka and the guys next door: Wouter, Mlawule and Michael van Leeuwen) and group members are always good for a discussion or some laughter. I thank in particular Xiaonan and Tony Zhou: the frequent table tennis sessions during my last year not only greatly improved my playing skills but at the same time protected me from being over-stressed. I also thank Xiaonan & Ma, Tony and his ‘assistant’ Boudewijn, for a great deal of ‘gezelligheid’ outside the lab. My thank also goes to my roommates (Zijian Tang and later Hetian Zhan), the Chinese friends in Delft, in particular previous and current board members of the Delft branch of ACSSNL, and members of the Fotoclub FC’59 (in particular Theo van der Heijdt), although the last two years I was unfortunately too busy to be truly active.

I have to acknowledge the architects at Spatial Phenomenon Atelier, Wenqian and Xiaoxi Hui, for designing the cover of this thesis. The cover image is a blurred and pixelized photo of yeast under the microscope, in which one could barely recognize the yeast cells. This might reflect the limited understanding of an architect of my research, but at the same time illustrates properly how little scientists understand this tiny microorganism.

It is impossible to name everyone who has supported me direct and indirectly, but you have to draw a line somewhere. I therefore would like to end the ‘Dankwoord’ with thanks to Zijian, Tobias Bartek, Horst Kiehl and Aljoscha Wahl, for providing tips for the lay-out of this thesis in L^AT_EX.

Liang Wu
Delft, September 2005