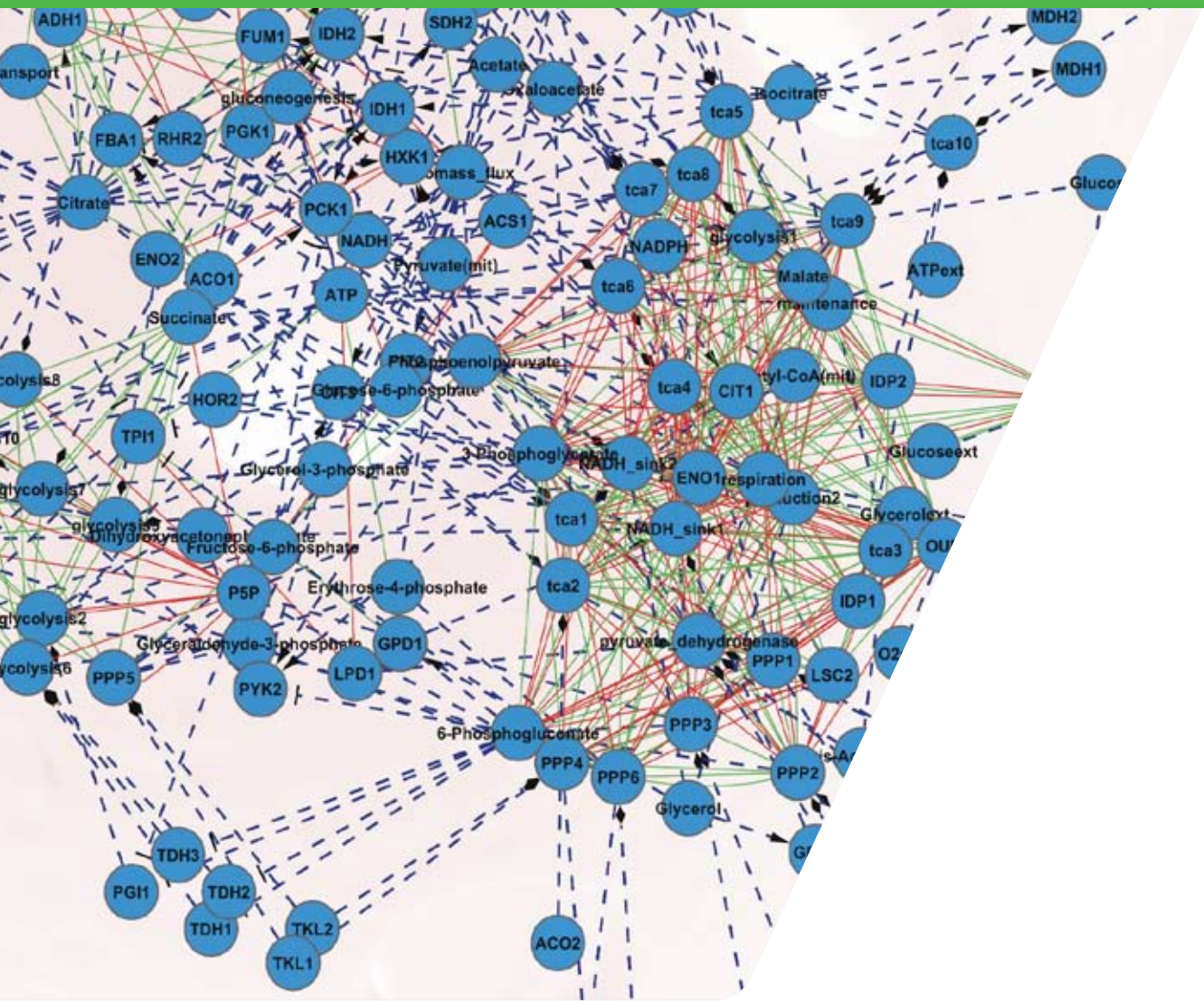

Paula Jouhten

Metabolic modelling and ^{13}C flux analysis

Application to biotechnologically important yeasts and a fungus

VTT PUBLICATIONS 724

Metabolic modelling and ^{13}C flux analysis

Application to biotechnologically important yeasts and a fungus

Paula Jouhten

Department of Biomedical Engineering and Computational Science
Faculty of Information and Natural Sciences
Helsinki University of Technology, Espoo, Finland

Graduate School in Computational Biology, Bioinformatics, and Biometry (ComBi)

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission of the Faculty of Information and Natural Sciences, for public examination and debate in Auditorium in the building of the Department of Applied Physics (Otakaari 3) at Helsinki University of Technology (Espoo, Finland) on the 3rd of December, 2009, at 12 noon.

ISBN 978-951-38-7371-4 (soft back ed.)
ISSN 1235-0621 (soft back ed.)

ISBN 978-951-38-7372-1 (URL: http://www.vtt.fi/publications/index.jsp)
ISSN 1455-0849 (URL: http://www.vtt.fi/publications/index.jsp)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT, Vuorimiehentie 5, PL 1000, 02044 VTT
puh. vaihde 020 722 111, faksi 020 722 4374

VTT, Bergsmansvägen 5, PB 1000, 02044 VTT
tel. växel 020 722 111, fax 020 722 4374

VTT Technical Research Centre of Finland, Vuorimiehentie 5, P.O. Box 1000, FI-02044 VTT, Finland
phone internat. +358 20 722 111, fax + 358 20 722 4374

Edita Prima Oy, Helsinki 2009

Paula Jouhten. Metabolic modelling and ^{13}C flux analysis. Application to biotechnologically important yeasts and a fungus [Aineenvaihdunnan mallinnus ja ^{13}C-vuoanalyysi. Sovellukset bioteknologisesti tärkeisiin hiivoihin ja homeeseen]. Espoo 2009. VTT Publications 724. 94 p. + app. 83 p.

Keywords metabolic modelling, metabolic flux, metabolic flux analysis (MFA), ^{13}C-labelling, ^{13}C-MFA, nuclear magnetic resonance (NMR) spectroscopy

Abstract

All bioconversions in cells derive from metabolism. Microbial metabolisms contain potential for bioconversions from simple source molecules to unlimited number of biochemicals and for degradation of even detrimental compounds. Metabolic fluxes are rates of consumption and production of compounds in metabolic reactions. Fluxes emerge as an ultimate phenotype of an organism from an integrated regulatory function of the underlying networks of complex and dynamic biochemical interactions. Since the fluxes are time-dependent, they have to be inferred from other, measurable, quantities by modelling and computational analysis. ^{13}C-labelling is crucial for quantitative analysis of fluxes through intracellular alternative pathways. Local flux ratio analysis utilises uniform ^{13}C-labelling experiments, where the carbon source contains a fraction of uniformly ^{13}C-labelled molecules. Carbon-carbon bonds are cleaved and formed in metabolic reactions depending on the *in vivo* fluxes. ^{13}C-labelling patterns of metabolites or macromolecule components can be detected by mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy. Local flux ratio analysis utilises directly the ^{13}C-labelling data and metabolic network models to solve ratios of converging fluxes.

In this thesis the local flux ratio analysis has been extended and applied to analysis of phenotypes of biotechnologically important yeasts *Saccharomyces cerevisiae* and *Pichia pastoris*, and a fungus *Trichoderma reesei*. Oxygen dependence of *in vivo* net flux distribution of *S. cerevisiae* was quantified by using local flux ratios as additional constraints to the stoichiometric model of the central carbon metabolism. The distribution of fluxes in the pyruvate branching point turned out to be most responsive to different oxygen availabilities. The distribution of fluxes was observed to vary not only between the fully respiratory, respiro-fermentative and fermentative metabolic states but also between different respiro-fermentative states. The local flux ratio analysis was extended

to the case of two-carbon source of glycerol and methanol co-utilisation by *P. pastoris*. The fraction of methanol in the carbon source did not have as profound effect on the distribution of fluxes as the growth rate. The effect of carbon catabolite repression (CCR) on fluxes of *T. reesei* was studied by reconstructing amino acid biosynthetic pathways and by performing local flux ratio analysis. *T. reesei* was observed to primarily utilise respiratory metabolism also in conditions of CCR. *T. reesei* metabolism was further studied and L-threo-3-deoxy-hexulosonate was identified as L-galactonate dehydratase reaction product by using NMR spectroscopy. L-galactonate dehydratase reaction is part of the fungal pathway for D-galacturonic acid catabolism.

Paula Jouhten. Metabolic modelling and ^{13}C flux analysis. Application to biotechnologically important yeasts and a fungus [Aineenvaihdunnan mallinnus ja ^{13}C-vuoanalyysi. Sovellukset bioteknologisesti tärkeisiin hiivoihin ja homeeseen]. Espoo 2009. VTT Publications 724. 94 s. + liitt. 83 s.

Avainsanat metabolic modelling, metabolic flux, metabolic flux analysis (MFA), ^{13}C-labelling, ^{13}C-MFA, nuclear magnetic resonance (NMR) spectroscopy

Tiivistelmä

Aineenvaihdunta kattaa kaikki biomuunnokset soluissa. Mikrobiaineenvaihdunta mahdollistaa yksinkertaisten lähtöaineiden muuntamisen rajoittamattomaksi määräksi erilaisia biokemikaaleja ja jopa haitallisten aineiden hajottamisen. Aineenvaihduntavuot ovat yhdisteiden kulutus- ja tuottonopeuksia aineenvaihdunnan reaktioissa. Vuot ilmentyvät organismin todellisena fenotyyppinä, jota säätelevät yhteistoiminnallisesti solun monimutkaiset ja dynaamiset vuorovaikutusverkot. Koska vuot ovat aikariippuvaisia, ne on määritettävä mallinnuksen ja laskennallisen analyysin avulla toisista, mitattavissa olevista, suureista. ^{13}C-leimaus on välttämätöntä, jotta vuot vaihtoehtoisilla solunsisäisillä reiteillä voidaan määrittää kvantitatiivisesti. Paikallisessa vuosuhdeanalyysissa käytetään tasaista ^{13}C-leimausta, jossa hiilenlähde sisältää osuuden täydellisesti ^{13}C-leimattuja molekyylejä. *In vivo* -vuot määräävät missä suhteissa aineenvaihdunnassa katkeaa ja muodostuu uusia hiili-hiilisidoksia. Aineenvaihdunnan välituotteiden ja makromolekyylien komponenttien ^{13}C-leimauskuvioita voidaan mitata massaspektrometrialla (MS) tai ydinmagneettisella resonanssispektroskopialla (NMR). Paikallisessa vuosuhdeanalyysissa käytetään suoraan mittausinformaatiota ^{13}C-leimauskuvioista ja aineenvaihduntaverkkomalleja vuosuhteiden ratkaisemiseksi.

Väitöskirjassa paikallista vuosuhdeanalyysia laajennettiin ja sovellettiin bioteknologisesti tärkeiden hiivojen *Saccharomyces cerevisiae* ja *Pichia pastoris*, ja homeen *Trichoderma reesei* fenotyyppien analysoimiseksi. *S. cerevisiae*n *in vivo* -vuojakauman riippuvuus hapen saatavuudesta määritettiin kvantitatiivisesti käyttämällä paikallisia vuosuhteita lisärajoitteina keskeisen hiiliaineenvaihdunnan stoikiometriselle mallille. Pyruvaattiristeyksen vuojakauma osoittautui herkimmäksi eri happisaatavuuksille. Selvästi erilaiset vuojakaumat havaittiin täysin respiratiivisessa, respiro-fermentatiivisessa ja täysin fermentatiivisessa aineenvaihdunnan tilassa, mutta myös eri respiro-fermentatiivisissa tiloissa. Paikallinen vuosuhdeanalyysi laajennettiin kahden hiilenlähteen tapaukseen, jossa

P. pastoris kulutti samanaikaisesti glyserolia ja metanolia. Metanolin osuudella kokonaishiilenlähteessä ei ollut yhtä merkittävää vaikutusta vuojakaumaan kuin hiivan kasvunopeudella. Hiilikataboliittirepression (CCR) vaikutusta *T. reesei*n vuojakaumaan tutkittiin rekonstruoimalla aminohapposynteesireitit ja tekemällä paikallinen vuosuhdeanalyysi. *T. reesei*n havaittiin käyttävän pääasiassa respiratiivista aineenvaihduntaa myös repressoivissa olosuhteissa. NMR-spektroskopiaa käytettiin myös D-galakturonihapon kaboliareitin tutkimuksessa ja L-treo-3-deoksi-heksulonaatti tunnistettiin *T. reesei*n L-galaktonaattidehydrataasireaktion tuotteeksi.

Preface

This study was mainly carried out at VTT Technical Research Centre of Finland in the team of Metabolic Engineering. In addition, a part of the work was done at the Department of Computer Science, University of Helsinki. The financial support from the Finnish Funding Agency for Technology and Innovation (Tekes), Academy of Finland (Finnish Centre of Excellence programme), Graduate School in Computational Biology, Bioinformatics, and Biometry (ComBi), and Finnish Foundation for Technology Promotion (TES), is appreciatively acknowledged.

Former Vice President Juha Ahvenainen, Vice President Anu Kaukovirta-Norja, Vice President Hans Söderlund, Technology Manager Richard Fagerström, and Technology Manager Tiina Nakari-Setälä are gratefully thanked for creating excellent working facilities and possibilities. Research Professor Merja Penttilä and Team Leader Laura Ruohonen at VTT are delightfully acknowledged for the scientific guidance and their positive attitude towards this thesis work. I am also grateful to Professor Esko Ukkonen at the Department of Computer Science, University of Helsinki and Professor Jens Nielsen at Chalmers University of Technology for providing excellent facilities and high level scientific guidance during my visits.

I wish to express my deepest gratitude to my supervisor docent Dr. Hannu Maaheimo. With patience he introduced me to the field of NMR spectroscopy, ^{13}C-labelling, and metabolic fluxes and since then he has always been there to listen to me and discuss my endless flux of ideas and to support and encourage me in all situations. Hannu's enthusiasm and interest in a wide range of topics in life never stop impressing me.

My sincere thanks go to Esa Pitkänen and Dr. Ari Rantanen for the collaboration, the willingness to share knowledge, and friendship from the days when we

did not share a language to present. In our collaboration I have been especially fond of the scientific enthusiasm and special kind of humour.

I would also like to acknowledge all other co-authors of my publications: Dr. Aina Sola from the group of Professor Pau Ferrer, Dr. Tiina Pakula and Team Leader Dr. Markku Saloheimo from the group of Protein Production at VTT, Eija Rintala, Satu Hilditch, Dr. Mervi Toivari and Dr. Marilyn Wiebe from the team of Metabolic Engineering at VTT, Professor Juho Rousu from the Department of Computer Science, University of Helsinki. In addition, Dr. Marilyn Wiebe and Vytas Raulinaitis are acknowledged for fruitful discussions and revising the English language of my manuscripts.

I would like to warmly thank all my colleagues at the NMR laboratory in Viikki, at VTT in Otaniemi, and at the Department of Computer Science in Kumpula for creating friendly working atmosphere and providing assistance whenever needed. It has been a pleasure and privilege to work with all of them! I wish to express special thanks to Dr. Kai Fredriksson for encouraging words in moments of self-doubt and for bewildering discussions. My warm thanks are also due to the people in the Nielsen's Lab for Systems Biology at Chalmers University of Technology for the joyful and enthusiastic environment and all collaboration during my visit and ever since. I would also like to express my cordial thanks to Dr. Kiran Patil from DTU Technical University of Denmark for the collaboration and friendship.

I gratefully thank Professor Risto Renkonen and Professor Ulla Ruotsalainen for critically reviewing the manuscript of this thesis.

An incredible source of energy and strength for this work has been my parallel life on floorball fields. I've been living my dreams while doing research and being an athlete. I wish to thank all the floorball people in Finland and abroad, who have brought so much joy to my life. Especially, I would like thank the people that I have got to know in Tapanilan Erä, Kassu, and all the coaches who have guided my sports career. To Erik Piispa I would like to express my deepest thanks for all invaluable guidance and encouragement. Rache and Fredrik I wish to thank for the special friendship during my visit in Sweden. My warmest thanks go to Harri for the precious friendship and never-ending encouragement.

Finally, I would like to express the greatest thanks to my mother Helena for her care, support and delicate guidance in my life. Without her I would have never had all the possibilities to succeed. To my father Mikko I wish to express my warmest gratitude for the support in different venues of my life and for lighting my enthusiasm for sports.

List of publications

I Jouhten, P., Rintala, E., Huuskonen, A., Tamminen, A., Toivari, M., Wiebe, M., Ruohonen, L., Penttilä, M. & Maaheimo, H. Oxygen dependence of metabolic fluxes and energy generation of *Saccharomyces cerevisiae* CEN.PK113-1A. *BMC Systems Biology* 2008, 2:60.

II Solà, A.*, Jouhten, P.*, Maaheimo, H., Sánchez-Ferrando, F., Szyperski, T. & Ferrer, P. Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates. *Microbiology* 2007, 153:281–290. *equal contributions

III Jouhten, P., Pitkänen, E., Pakula, T., Saloheimo, M., Penttilä, M. & Maaheimo, H. ^{13}C-metabolic flux ratio and novel carbon path analyses confirmed that *Trichoderma reesei* uses primarily the respirative pathway also on the preferred carbon source glucose. *BMC Systems Biology* 2009, 3:104.

IV Rantanen, A., Rousu, J., Jouhten, P., Zamboni, N., Maaheimo, H. & Ukkonen, E. An analytic and systematic framework for estimating metabolic flux ratios from ^{13}C tracer experiments. *BMC Bioinformatics* 2008, 9:266.

V Kuorelahti, S., Jouhten, P., Maaheimo, H., Penttilä, M. & Richard, P. L-galactonate dehydratase is part of the fungal path for D-galacturonic acid catabolism. *Molecular Microbiology* 2006, 61:1060–1068.

Author’s contributions

In Publication I the author of the thesis participated in the design of the study, performed chemostat cultivations, carried out the nuclear magnetic resonance (NMR) spectroscopic experiments, designed and performed the modeling and the computational work and wrote the manuscript. In Publication II the author of the thesis performed NMR spectroscopic experiments, carried out the analysis of the NMR spectral data and participated in the preparation of the manuscript. In Publication III the author of the thesis participated in the design of the study, performed the cultivations, carried out the NMR spectroscopic experiments and performed the ^{13}C-metabolic flux ratio analysis and analysed the results. The author of the thesis also together with the second author interpreted the results of the computational pathway analysis and wrote the manuscript. In Publication IV the author of the thesis constructed the models of the metabolic networks and contributed to the development of the computational methods. In Publication V the author of the thesis performed NMR spectroscopic experiments and analysis of NMR spectral data.

Contents

Abstract 3

Tiivistelmä 5

Preface 7

List of publications 9

Author's contributions 10

Abbreviations 13

1. Introduction – biology part 17
   - 1.1 Cell factories and model organisms 17
   - 1.2 Metabolism 18
   - 1.3 Metabolic fluxes 19
   - 1.4 Regulation of flux phenotype 19
   - 1.5 Oxygen affects flux phenotype 20
     - 1.5.1 Oxygen responsive hierarchical regulatory mechanisms 22
   - 1.6 Carbon catabolite repression regulation of phenotype 23

2. Introduction – method part 25
   - 2.1 Systems biology 25
   - 2.2 Metabolic modelling for flux analysis 25
     - 2.2.1 Stoichiometric models 26
       - 2.2.1.1 Genome-wide metabolic reconstructions 26
     - 2.2.2 Kinetic models 27
   - 2.3 Metabolic flux analysis 28
     - 2.3.1 Constraint-based analysis 30
     - 2.3.2 ^{13}C-metabolic flux analysis 31
       - 2.3.2.1 ^{13}C-labelling and analytical methods 32
       - 2.3.2.2 ^{13}C-metabolic flux analysis – mathematical and statistical methods 36
       - 2.3.2.3 ^{13}C-metabolic flux analysis in large scale networks 41

3. Aims of the research 43
   - 3.1 Oxygen dependence of fluxes and underlying regulation in *S. cerevisiae* 43

3.2 Two carbon source case of methanol and glycerol utilisation by *P. pastoris* 44
3.3 Path identification and the effect of carbon catabolite repression on metabolic fluxes in *T. reesei* 44
3.4 Framework for analytical determination of flux ratios 45
3.5 NMR spectroscopy as a tool in pathway identification 46

4. Research methods 47
4.1 Strains 47
4.2 Cultivations 47
4.3 Biosynthetically directed fractional ^{13}C-labelling 48
4.4 Sampling 50
4.5 NMR spectroscopy 50
4.6 Metabolic flux ratio analysis 51
4.7 Metabolic modelling for ^{13}C-metabolic flux analysis 53
4.8 Pathway reconstruction 59
4.9 Localization of amino acid biosynthetic enzymes in *T. reesei* 61

5. Results and discussion 62
5.1 Utilization of ^{13}C-metabolic flux analysis excluded cofactor mass balances 63
5.2 Pyruvate branching point distribution responsive 63
5.3 Methanol and glycerol co-utilization extension 64
5.4 Flux distributions robust against different fractions of methanol 65
5.5 Metabolic flux ratio analysis of *T. reesei* necessitated reconstruction of biosynthetic pathways of amino acids 66
5.6 Primary respiratory metabolism 66
5.7 Framework for analytic and systematic derivation of flux ratio equations 68
5.8 L-threo-3-deoxyhexulosonate is a reaction product of L-galactonate dehydratase 69

6. Conclusions and prospects 71
6.1 Robust regulatory system enables stable flux phenotype 71
6.2 Determinants of energy generation processes 74
6.3 Prospects of local flux ratio analysis 75
6.4 Large-scale flux analysis 76
6.5 Flux analysis in dynamic conditions 76

References 78

Appendices

Publications I–V

Appendix V of this publication is not included in the PDF version. Please order the printed version to get the complete publication (http://www.vtt.fi/publications/index.jsp).

Abbreviations

AcCoA	acetyl coenzyme A
Ala	L-Alanine
Arg	L-Arginine
Asp	L-Aspartic acid
ATP	adenosine 5-triphosphate
BDF	biosynthetically directed fractional
CCR	carbon catabolite repression
CDW	cell dry weight
CER	carbon dioxide evolution rate
CE-TOFMS	capillary electrophoresis-time-of-flight-mass spectrometry
COSY	correlation spectroscopy
D	dilution rate
DHAP	dihydroxyacetonephosphate
D_2O	deuterium oxide
EMU	elementary metabolite unit
E4P	D-erythrose 4-phosphate
FBA	flux balance analysis
FID	free induction decay
GC-MS	gas chromatography-mass spectrometry
Glu	L-Glutamic acid

Gly	Glycine
His	L-Histidine
HSQC	heteronuclear single quantum correlation
Ile	L-Isoleucine
LC-MS	liquid chromatography-mass spectrometry
LC-MS/MS	liquid chromatography-tandem mass spectrometry
Leu	L-Leucine
Lys	L-Lysine
MAE	malic enzyme
METAFoR	metabolic flux ratio
MFA	metabolic flux analysis
MOMA	minimisation of metabolic adjustment
mRNA	messenger ribonucleic acid
MS	mass spectrometry
NAD^+	nicotinamide adenine dinucleotide (oxidized)
NADH	nicotinamide adenine dinucleotide (reduced)
$NADP^+$	nicotinamide adenine dinucleotide phosphate (oxidized)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced)
NMR	nuclear magnetic resonance
NO	nitric oxide
Oaa	oxaloacetate
Oaa_{cyt}	cytosolic oxaloacetate
Oaa_{mit}	mitochondrial oxaloacetate
Oga	oxoglutarate
OUR	oxygen uptake rate
PDH	pyruvate dehydrogenase
Pep	phosphoenolpyruvate
PEPck	phosphoenolpyruvate carboxykinase

Phe	L-Phenylalanine
PPP	pentose phosphate pathway
Pro	L-Proline
Pyr	pyruvate
Pyr_{cyt}	cytosolic pyruvate
Pyr_{mit}	mitochondrial pyruvate
ROS	reactive oxygen species
R5P	D-ribose 5-phosphate
sd	standard deviation
SEM	standard error of the mean
Ser	L-Serine
S7P	D-sedoheptulose 7-phosphate
TCA	tricarboxylic acid
Thr	L-Threonine
TOCSY	total correlation spectroscopy
TOF	time-of-flight
Tyr	L-Tyrosine
T3P	triose 3-phosphates
YNB	yeast nitrogen base

1. Introduction – biology part

1.1 Cell factories and model organisms

Biotechnology offers possibilities for production of highly specialised biochemicals and for sustainable and economic process alternatives. It exploits cells or biocatalytes, enzymes, in bioconversions. Cells contain complete machineries for bioconversions from simple source molecules to unlimited number of biochemicals and degradation potential for breakdown of compounds even hazardous. Microorganisms are efficient cell factories whose requirements on the process conditions are modest and they often grow on inexpensive media. Even waste streams or effluents can be utilised as raw materials for bioprocesses.

Eukaryotic microorganism, yeast *Saccharomyces cerevisiae* (Figure 1), has a long history of biotechnological utilisation from conventional use as baker's yeast to production of diverse biochemicals. Concomitantly *S. cerevisiae* has been widely applied as a model organism in studies of general cell physiology. Due to the broad interest and the long history, *S. cerevisiae* is one of the most studied microorganisms with highly developed molecular biology tools and modelling frameworks [Petranovic and Vemuri, 2009; Herrgård *et al.*, 2008; Nevoigt, 2008]. Today *S. cerevisiae* is emerging as a simple eukaryotic model, a systems biology workhorse, for elucidating the mechanisms of even human diseases [Petranovic and Nielsen, 2008; Chen and Thorner, 2007]. The focus in systems biology is in understanding the function of a cell system as a whole [Lazebnik, 2002]. Since the complexity of cell systems is beyond intuitive comprehension, the core of systems biology is mathematical modelling of biological processes [Kitano, 2002]. The significant similarity of the cell function among eukaryotic cells offers promising prospects for *S. cerevisiae* models [Chen and Thorner, 2007; Petranovic and Nielsen, 2008; Botstein *et al.*, 1997]. Physiology of *S. cerevisiae* was studied in Publication I of the thesis. The organisms investi-

gated in Publications II, III and V, are fungus *Trichoderma reesei* (Figure 1) and yeast *Pichia pastoris* that are two host organisms for industrial production of natural and heterologous proteins.

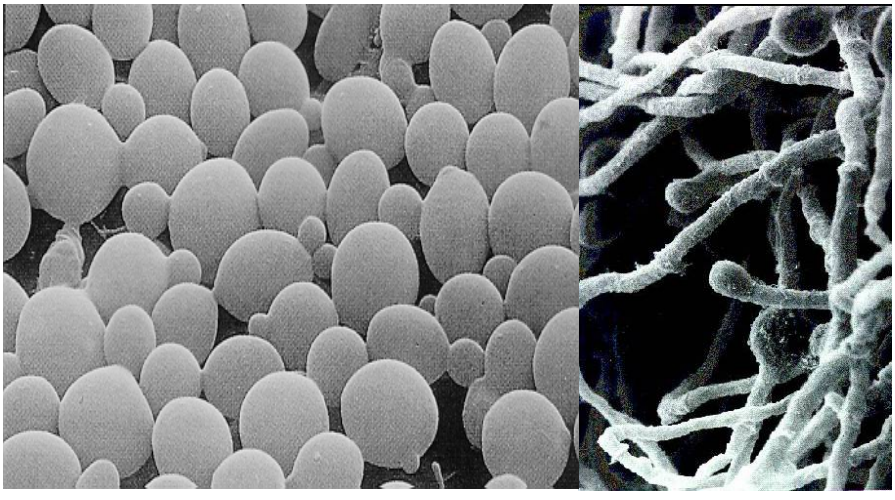

Figure 1. Budding yeast *S. cerevisiae* (on the left) and filamentous fungus *T. reesei* (on the right) are biotechnologically important production organisms.

1.2 Metabolism

All bioconversions in cells derive from metabolism. Metabolism is a set of biochemical reactions made feasible by enzymes [Stryer, 1995], which in turn are encoded by genes. Thus, the whole potential of metabolism of an organism is encoded in its genome, the complement of all genes. At present, the metabolic potential of an organism can usually be efficiently modified or engineered, with the variety of molecular biology tools available today. Metabolic engineering is, as stated by Stephanopoulos already in 1990's, "directed improvement of product formation or cellular properties through the modification of specific biochemical reaction(s) or the introduction of new one(s) with the use of recombinant DNA technology" [Stephanopoulos *et al.*, 1998].

Unicellular organisms comprise catabolism of substrates and anabolic pathways for synthesis of biomass components in a single cell. Pathways for metabolization of various carbon sources unite and a range of biosynthetic pathways initiate in central carbon metabolism, which is a common knot for catabolism and anabolism [Ma and Zeng, 2003]. Source molecules are broken down and

energy and precursors for biosynthetic pathways are produced in the central carbon metabolism.

1.3 Metabolic fluxes

Metabolic flux is a time-dependent quantity of the rate of consumption or production of compounds in a metabolic reaction [Stephanopoulos *et al.*, 1998; Nielsen *et al.*, 2003]. Metabolic fluxes are process streams of a cell factory. Therefore, a common aim of metabolic engineering is to generate changes in pathway fluxes. The essential biocatalytes, i.e. enzymes, can be amplified, deleted, and modified with versatile molecular biology tools. However, metabolic homeostasis prevails in cells that are highly balanced systems. A flux through a pathway depends on various factors in addition to the enzymes. The complement of fluxes in a cell, namely fluxome is cell's ultimate response to genetic and environmental conditions [Sauer, 2004]. The flux response emerges from an integrated function of complex and dynamic interaction networks (metabolic, signal transduction, regulatory, protein-protein interaction networks etc). Many of the components of the biochemical interaction networks such as concentrations of enzymes, other proteins, metabolites, and genome-wide gene expression levels are at present directly measurable with high-throughput systems. Since the fluxes are dependent of time, they cannot be directly measured but have to be inferred from other, measurable, quantities through a model based computational analysis.

1.4 Regulation of flux phenotype

The intertwined biochemical interaction networks of a cell form a regulatory system. The complex regulatory system enables both fine-tuned adaptive responses and robustness of the phenotype against genetic defects and fluctuations in external conditions [Kitano, 2007]. The ability to adapt to the prevailing growth conditions is essential for micro-organisms like *S. cerevisiae* that are unable to control the extracellular conditions. Furthermore, the regulatory system is capable of attenuating effects of genetic modifications on phenotype [Davies and Brindle, 1992; Schaaff *et al.*, 2004; Blank *et al.*, 2005].

The regulation of a finite change in flux can be conveniently and quantitatively divided into hierarchical and metabolic regulation [ter Kuile and Westerhoff, 2001]. Hierarchical regulation covers the steps of the central dogma of

molecular biology: gene expression, transcription, and translation. It ultimately determines the amounts of enzymes. Transcription is regulated by transcription factor proteins and other regulatory factors that may bind specifically to a gene to initiate or to speed up the formation of a messenger RNA (mRNA) [Fuda *et al.*, 2009]. The mRNA is then transferred out of the nucleus into the cytosol where it binds to ribosomes for translation. Rate of translation is dependent on various factors including ribosome density [Arava *et al.*, 2003; Brockmann *et al.*, 2007]. The degradation rates also affect the quantities of mRNAs and proteins.

Metabolic regulation of a change in flux includes everything beyond the enzyme concentrations such as activation of enzymes and kinetic control of reactions [ter Kuile and Westerhoff, 2001]. Post-translational modifications of proteins modulate their activity [Uy and Wold, 1977; Mann and Jensen, 2003]. For example phosphorylation can fully determine the activity of an enzyme [Ptacek *et al.*, 2005]. Signal transduction cascades pass phosphorylations as a response of sensing the growth conditions [Zaman *et al.*, 2008]. Reaction rates depend on the concentrations of the reactants after the particular kinetics of an enzyme [Stryer 1995]. In addition, metabolites can act as allosteric effectors and affect the reaction rates [Monod *et al.*, 1965] or even trigger regulation on the hierarchical regulatory levels [Sellick and Reece, 2003].

Metabolic homeostasis derives from dependences between flux, enzymes and metabolites and interdependences between reactions created by the metabolic network. Furthermore, hubs [Ma'ayan, 2009] such as cofactors NADH and NADPH and the energy unit ATP, are common metabolites for the whole network and create regulatory dependences even between distant pathways of the metabolic network. While there is lack of detailed knowledge on kinetic parameters and reaction mechanisms of large fraction of metabolic enzymes, thermodynamics provides insight to the dependences between fluxes and metabolites [Kummel *et al.*, 2006; Beard and Qian, 2005].

1.5 Oxygen affects flux phenotype

Oxygen conditions in nature vary between the oxygen partial pressure in air and complete anaerobiosis. Microorganisms have adapted to different ranges of oxygen availabilities depending on their natural habitats. A central role of oxygen metabolism is highly usual for any biological system [Koch and Britton, 2008]. Aerobic organisms are able to utilise oxygen as the final electron acceptor in the electron transfer chain, which is coupled to ATP synthesis. Since oxygen has

high electronegativity, the electron transfer reactions provide large transfer energy and the respiratory ATP production has an extremely high yield [Koch and Britton, 2008]. On the other hand, aerobic organisms need protection mechanisms against the deteriorating effects of oxygen [Jamieson, 1998]. Cell components are deteriorated by external oxidants and oxygen radicals generated in cell's internal oxygen utilising processes [Herrero *et al.*, 2008]. In should be noted that oxidation severely damages proteins, lipids, and nucleic acids [Jamieson, 1998; Herrero *et al.*, 2008]. Oxidative stress signalling activates repair mechanisms and degradation pathways for damaged components [Letavayová *et al.*, 2006]. Interestingly, cells counteract oxidative and reductive stresses with at least partly overlapping mechanisms [Trotter and Grant, 2002].

Aerobic environment poses a challenge also on the regulation of cell's redox balance. Cell cytosol is normally maintained reductive [López-Mirabal and Winther, 2008] and glutathione is the main buffer in the redox balancing system. In addition, it is linked to numerous cellular processes like membrane transport systems and carbon and nitrogen metabolisms [Perrone *et al.*, 2005; López-Mirabal and Winther, 2008]. Balanced redox conditions affect the metabolic homeostasis also because several metabolic reactions are redox reactions. When oxygen is not available as an electron acceptor, *S. cerevisiae* produces glycerol as a redox sink [Bakker *et al.*, 2001]. Fermentative pathway is redox neutral, but glycerol production occurs when the amount of NADH formed in biosynthesis exceeds the capacity of respiration to regenerate NADH to NAD^+. Reoxidation of NADH is prioritised under the conditions of oxygen deficiency and the carbon flux is directed to the fermentative pathway instead of to the TCA cycle [Weusthuis *et al.*, 1994; Publication I].

In *S. cerevisiae* oxygen limitation in the extracellular medium shifts the flux phenotype. The glycolytic flux is increased and ethanol production takes place [Weusthuis *et al.*, 1994; Publication I]. In the absence of ethanol production, metabolism is fully respirative. Respiro-fermentative phenotypes are observed in conditions of limited respiration. Respiratory limitation faces *S. cerevisiae* not only in lack of oxygen but also under excess glucose conditions and at high growth rate [Cortassa and Aon, 1998]. Glucose sensing and signalling network is active in conditions of excess glucose and it represses the components of the respiratory chain and the TCA cycle [Zaman *et al.*, 2009]. In contrast, the respiratory chain components have been observed to be upregulated under low oxygen conditions [Rintala *et al.*, 2009]. Aerobic alcoholic fermentation is observed in *S. cerevisiae* also at high growth rates when the glycolytic flux exceeds the

critical limit that depends on maximum respiratory rate [Vemuri *et al.*, 2007]. Exceeding the critical limit results in overflow metabolism in pyruvate branching point of central carbon metabolism and thus, to a flux to fermentative pathway [Vemuri *et al.*, 2007; Frick and Wittmann, 2005]. These observations indicate that similar metabolic states are generated by different transcriptional regulatory patterns highlighting the importance of the post-transcriptional and metabolic regulation of the phenotypes. Accordingly, the fluxes through glycolytic enzymes have previously been shown to be mainly regulated at post-transcriptional level [Daran-Lapujade *et al.*, 2007]. Furthermore, pure metabolic regulation can increase the glycolytic flux at least eight fold [van der Brink *et al.*, 2008] and the increased glycolytic flux observed under high temperature is primarily maintained by metabolic regulation [Postmus *et al.*, 2008]. On the other hand, the gluconeogenetic and glyoxylate cycle enzymes have been observed to be regulated at transcriptional level [Kolkman *et al.*, 2005].

Publication I studied the response of the metabolism of *S. cerevisiae* to the different oxygen provisions at the flux phenotypic level.

1.5.1 Oxygen responsive hierarchical regulatory mechanisms

The most well known oxygen-responsive hierarchical regulatory systems are dependent on the levels of heme and sterols [Hon *et al.*, 2003; Davies and Rine, 2006; Kwast *et al.*, 1998]. The synthesis of both of them requires molecular oxygen and thus, their levels decline in the depletion of oxygen. Hap-transcription factors respond to the levels of heme. Hap1 regulates the expression of anaerobic genes whereas Hap2/3/4/5 factors regulate expression of aerobic genes [Kwast *et al.*, 1998]. Hap1 factor has been shown to have a gentle slope in the activity profile in mild oxygen limitation but a sharp increase in severe lack of oxygen [Hon *et al.*, 2003]. Hap2/3/4/5 factors regulate genes encoding metabolic TCA cycle enzymes among others and Hap4 particularly activates catabolism of respiratory carbon sources like ethanol [Raghevendran *et al.*, 2006].

The mitochondrial production of oxidative stress mediating reactive oxygen species (ROS) and nitric oxide (NO) in low oxygen conditions have been suggested to be involved in signalling for induction of hypoxic genes [Castello *et al.*, 2006; D'Autréaux and Toledano, 2007; Woo *et al.*, 2009]. Accordingly transient oxidative stress response has been observed in *S. cerevisiae* in sudden depletion of oxygen [Dirmeier *et al.*, 2002]. Genes encoding enzymes involved in biosynthesis of fatty acids, which requires oxygen, belong to hypoxic genes.

The mitochondria are also known to signal of a respiratory defect by retrograde signalling that affects even the transcription of nuclear genes [Butow and Avadhani, 2004; Liu and Butow, 2006]. Retrograde signalling coordinates carbon and nitrogen metabolisms to respond to the requirements of the deficient state. In conditions of mitochondrial deficiency, the regulation of genes encoding TCA cycle enzymes switch from Hap-complex to retrograde regulators [Liu and Butow, 2006].

1.6 Carbon catabolite repression regulation of phenotype

Carbon catabolite repression (CCR) is a phenomenon where in presence of a preferred carbon source the pathways for utilisation of alternative carbon sources are repressed [Gancedo, 1998]. In the presence of excess glucose CCR of *S. cerevisiae* strongly represses metabolization of other carbon sources and also the respirative pathway as discussed above [Gancedo 1998; Zaman *et al.*, 2009; Westergaard *et al*., 2007]. High glucose mediates the redistribution of fluxes to respirative and fermentative pathways similarly as varying oxygen conditions and exceeding the maximum respiratory capacity [Gombert *et al.*, 2001; Nissen *et al.*, 1997; Vemuri *et al.*, 2007; Publication I]. Fermentation and high glycolytic flux enable high rate of ATP production. Glucose repression in *S. cerevisiae* is a regulatory switch that prefers high rate of ATP production instead of the high ATP yield that could be obtained from the respirative pathway.

T. reesei is naturally adapted to grow in nutrient poor environments, where it is able to use complex plant material as carbon source. *T. reesei* and number of other filamentous fungi and cellulolytic bacteria produce and secrete plant polymer hydrolyzing enzymes such as cellulases and hemicellulases to their surroundings to break down the polymers into easily metabolizable monomers [Kumar *et al.*, 2008]. The powerful machinery producing hydrolytic enzymes in *T. reesei* is under CCR when a preferred carbon source, such as glucose, is available. Small oligosaccharides or derivative parts of the polymers in the environment of the fungus act as inducers of expression of genes encoding hydrolytic enzymes. The inductive signaling is specific for particular sets of enzymes [Ilmén, 1997; Aro *et al*., 2005]. However, under high glucose concentrations, CCR overrules the inductive signals [Ilmén *et al*., 1997]. The regulatory switch of energy generation in *T. reesei* is different from the switch in *S. cerevisiae*. In *T. reesei* CCR does not cause repression of genes encoding the TCA cycle enzymes or the respiratory pathway components. Thus, CCR does not hinder the

high yield respirative energy generation in *T. reesei* [Chambergo *et al*., 2002; Gancedo, 1998]. The difference is reasonable in evolutionary sense because all the available energy in the nutritionally harsh natural habitats of *T. reesei* is valuable.

The signalling pathways for glucose repression in *S. cerevisiae* are widely studied [Zaman *et al*., 2009]. The systems biology approach has provided further understanding of the interaction of separate signalling pathways in *S. cerevisiae* in yielding specific responses to the growth conditions [Westergaard *et al*., 2007]. In *T. reesei* Cre1 is the key mediator protein of CCR [Strauss *et al.*, 1995; Ilmén *et al.*, 1996]. It is structurally highly similar to Mig1, a key protein in glucose repression in *S. cerevisiae*. Despite the sequence and structural similarity, the functional dissimilarities of Cre1 and Mig1 have led to the conclusion that glucose repression functionalities in filamentous fungi and yeasts have evolved separately [Cziferszky *et al.*, 2002; Vautard *et al*., 1999]. Pfeiffer *et al.* (2001) has also argued that the evolution from unicellular to undifferentiated multicellular organisms like *T. reesei* has been facilitated by the preference of high yield energy generation by respiration. The role of respirative metabolism in the development of multicellular organisms has recently been supported by Koch and Britton (2008).

In Publication III the distribution of intracellular metabolic fluxes in *T. reesei* were studied in different conditions of CCR.

2. Introduction – method part

2.1 Systems biology

Systems biology focuses on system level function of cells instead of the conventional approach of mere concentration on individual components [Kitano, 2002; Lazebnik, 2002]. Thus, mathematical modelling, networks of biochemical interactions, and high-throughput methods for simultaneous profiling of large numbers of cell components are essence of systems biology. Modelling is crucial for studying highly complex biological systems. Cell components and processes transfer information through interactions which enables for example adaptation mechanisms for survival and, on the contrary, phenotypic robustness against fluctuations in environmental conditions [Kitano, 2007]. An ultimate aim of systems biology is to generate predictive *in silico* models of biological systems. Modelling is an iterative process of continuous improvement of the description of the system. Models are mathematical representations of phenomena of interest and they are always simplifications of the actuality. According to the retelling of Einstein´s statement: models should be as simple as possible, but not simpler, thus the level of simplification as well as the type of the model should be designed for the purpose of the model [Klipp, 2007].

2.2 Metabolic modelling for flux analysis

The simplest models of metabolism are *black box* models in which everything else than the external fluxes of uptake and secretion is hidden in the black box. Intracellular reactions are not specified but just wrapped into the box. An overall reaction equation describes the conversion of substrates to products in the black box. Despite the obvious simplicity, the black box models can be utilised for calculation of mass balances, elemental balances and degree of reduction bal-

ances in modelling of cell factories, thus, also in calculation of process figures like yields and productivities. However, to be able to engineer the process figures, information on what occurs inside the black box is of importance.

2.2.1 Stoichiometric models

Stoichiometric models of metabolism specify the individual reactions in the system and the reaction stoichiometry relations of substrates and products. Stoichiometric models are static models and thus, do not include any reaction kinetics. While kinetic models of any medium size branching networks are still inconvenient due to the lack of knowledge and computational challenges, stoichiometric models have proven to be highly useful in metabolic studies of large networks.

2.2.1.1 Genome-wide metabolic reconstructions

The emergence of efficient sequencing and DNA techniques brought along a growing number of published fully sequenced and annotated genomes of organisms. Even the complete metabolic potential of an organism is encoded in its genome. Therefore, the availability of the genomes and the development of comparative genomics lead into reconstruction of genome wide metabolic models. The first genome-wide metabolic network reconstruction of *S. cerevisiae* was done by Förster *et al.* (2003) (Figure 2). Automatic methods exist both for full reconstruction of genome wide metabolism and for pathway searches from the given substrate to a product [Feist *et al.*, 2009 (review); Karp *et al*., 2002; Pinney *et al.*, 2005; Notebaart *et al.*, 2006]. However, after the automatic work, reliable metabolic reconstruction requires laborious manual curation including both literature checks and experimental verification of the metabolic network model [Francke *et al.*, 2005; Feist *et al.*, 2009 (review); Herrgård *et al.*, 2008 (*S. cerevisiae* consensus model); Duarte *et al.*, 2007 (human); Shinfuku *et al.*, 2009 (*Corynebacterium glutamicum*)]. The characterisation of enzymes and verification of their products as was done in Publication V in the thesis contributes to experimental validation of models. The genome wide metabolic reconstructions are stoichiometric models including static reaction descriptions and preferentially annotations of enzymes catalysing the reactions. The genome-wide metabolic models offer frameworks for investigations of the complete metabolic potential of an organism, and for data interpretation and analysis [Patil and Nielsen, 2005].

They also provide scaffolds for models of smaller and dynamic systems, and importantly they provide a link between the genome and the metabolic reactions.

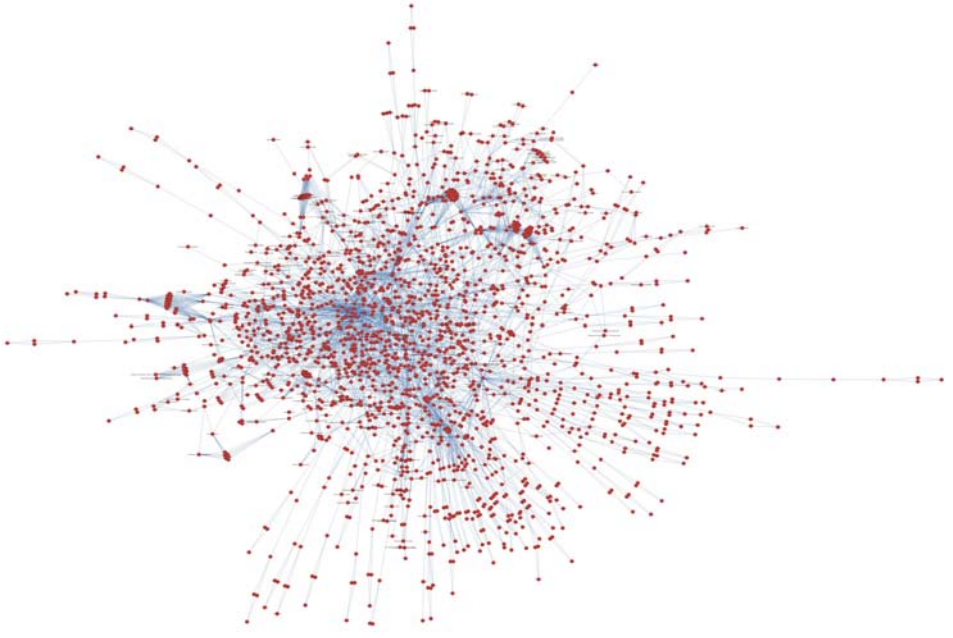

Figure 2. The genome-wide metabolic network of *S. cerevisiae* as a bipartite graph of metabolite and reaction nodes. The first genome-wide metabolic network reconstruction of *S. cerevisiae* reconstructed by Förster *et al.* (2003) included two compartments, cytosol and mitochondria, and 1175 metabolic reactions and 584 metabolites. The latest consensus model is divided into eight compartments and contains a total of 1168 metabolites and 1857 reactions and also 832 genes, 888 proteins, and 96 catalytic complexes [Herrgård *et al.*, 2008].

The reaction lists are readily converted into metabolite mass balances, functions of reaction rates, and further into stoichiometric matrices. Stoichiometric matrices can be analyzed by techniques of linear algebra to understand the metabolic potential and the structure of the metabolism of the particular organism [Palsson, 2006]. Stoichiometric reaction descriptions and a stoichiometric matrix are also the basic requirements for analysis of metabolic fluxes.

2.2.2 Kinetic models

Kinetic metabolic models include time-dependent mechanistic descriptions of reactions [Klipp, 2007]. Metabolic enzymes possess different mechanisms and

thus, the kinetic equations are enzyme specific [Stryer, 1995]. Depending on the reaction mechanism of an enzyme, the number of effectors and parameters varies. It is convenient to experimentally determine kinetics of isolated enzymes *in vitro*. However, it is likely that *in vivo* under crowded and compartmentalised conditions of a cell, the kinetics differ significantly from what is determined *in vitro*. Teusink *et al.* (2000) performed a study were a kinetic model of yeast glycolysis was set up and *in vitro* determined values for kinetic parameters were utilised for simulation. The output data from model simulations were compared to the experimental flux and metabolite data. There one finds significant discrepancies between simulated and experimental values of half of the reactions even though glycolysis is an extensively studied part of metabolism. Experimental data for computational estimation of kinetic parameters is acquired in perturbation experiments of the system of interest [Nikerel *et al.*, 2006; Vaseghi *et al.*, 1999]. However, since the number of parameters in kinetic metabolic models is usually huge, their estimation is demanding. Thus, different approximations of kinetic equations have been developed for studying the dynamic behaviour of a metabolic system [Heijnen, 2005 (review); Visser and Heijnen, 2003; Visser *et al.*, 2004; Savageau, 1970; Liebermeister and Klipp, 2006].

2.3 Metabolic flux analysis

Determination of intracellular *in vivo* fluxes is called metabolic flux analysis (MFA), which applies mass balances around metabolites according to the stoichiometric model. Thus, mass balances around metabolites are formulated as functions of the fluxes. Dilution of metabolite pools due to growth can be taken into account as a dilution term although in many cases the dilution term is negligible [Nielsen *et al.*, 2003].

$$\frac{dc_{met}}{dt} = v_{met} - \mu c_{met} \qquad (1)$$

In equation 1 c_{met} is the concentration of a particular metabolite, v_{met} the net rate of formation and consumption of the metabolite by all the fluxes in the system, and in the dilution term μ is the growth rate.

Integration of equation 1 over time for all the metabolites would yield the time dependence of concentrations $c_{met}(t)$. However, the fluxes v_{met} are often unknown functions of metabolite concentrations and unknown enzyme kinetic and other parameters as discussed above. Since it is challenging to simulate the dynamic

concentration profiles, steady state conditions are usually considered. In steady state conditions the growth rate and the metabolic fluxes are constant and there is no accumulation of intracellular metabolites but their pools also remain constant [Nielsen *et al.*, 2003]. Steady state microbial growth is reached in continuous cultures (chemostat cultures) where the rates of medium flow into and out from the bioreactor are equal. Pseudo-steady state conditions prevail in early exponential phase of a batch culture when the cells grow at maximum rate while the changes in the culture medium are still insignificant. Under assumption of steady state conditions the following form of equation 1 holds [Nielsen *et al.*, 2003]:

$$v_{met} - \mu c_{met} = 0 \tag{2}$$

Because the intracellular metabolite pools are generally very small, the dilution term is tiny compared to the fluxes producing and consuming metabolites, particularly in the central carbon metabolism that encompasses all the major fluxes [Stephanopoulos *et al.*, 1998]. When the dilution term can be assumed to be negligible, the equation 2 adapts a simple form:

$$v_{met} = 0 \tag{3}$$

which in matrix notation reads:

$$\mathbf{Nv} = \mathbf{0} \tag{4}$$

where **N** is the stoichiometric matrix with the fluxes of the system in columns and the stoichiometric coefficients of metabolites in each of the fluxes in rows and **v** is a column vector of fluxes. The stoichiometric matrix transforms the biology of metabolic reactions into mathematical framework. The matrix equation actually represents K linear mass balances for the K metabolites that contribute to J fluxes. Since there are always less metabolites K than fluxes J, the degree of freedom $F = K - J$ remains and some of the fluxes in vector v need to be determined to solve the rest of them [Nielsen *et al.*, 2003]. If the stoichiometric matrix is partitioned into two parts for measured fluxes ($\mathbf{N}_m$) and for unknown fluxes ($\mathbf{N}_u$), the equation can be rewritten in the following way:

$$\mathbf{Nv} = \mathbf{N}_m\mathbf{v}_m + \mathbf{N}_u\mathbf{v}_u = \mathbf{0} \tag{5}$$

where $\mathbf{v}_m$ is a vector of measured rates and $\mathbf{v}_u$ a vector of unknown rates. If exactly F fluxes have been measured and if $\mathbf{N}_u$ can be inverted, the unknown rates can be directly solved with matrix algebra [Stephanopoulos *et al.*, 1998]. If $\mathbf{N}_u$ has full rank ($rank(\mathbf{N}_u) = K$), it can be inverted and the unknown fluxes calcu-

lated by Gaussian elimination but if the rank of $\mathbf{N}_u$ is less than K, $\mathbf{N}_u$ is singular and the system is underdetermined. If the set of reaction stoichiometries are linearly dependent, the rank of $\mathbf{N}_u$ is less than K even though the number of measured fluxes equals degrees of freedom. Also if some of the measured rates are redundant, the rank of $\mathbf{N}_u$ is less than K.

Very seldom it is practically possible to measure enough fluxes to reach an algebraic solution to the metabolic system of linear mass balance equations. This is obviously always the case with large, genome-scale metabolic models.

2.3.1 Constraint-based analysis

The space of metabolic states has as many dimensions as reactions in the system. The stoichiometry of the reactions, equation 4, limits the space into a subspace that is a hyperplane. If the reactions are defined so that they are all positive, the plane is converted into a cone. If additionally upper bounds, maximum capacities, can be defined for the fluxes, a closed convex cone solution space is obtained. All the possible metabolic states of an organism, the feasible flux distributions, lie in that solution space. Thus, it is the space of phenotypes which an organism can express. To further shrink the solution space, additional constraints have been set up from reaction thermodynamics [Beard *et al.*, 2002; Beard *et al.*, 2004; Price *et al.*, 2004b; Price *et al.*, 2006], from experimental transcription data and from extracellular metabolome for condition-specific solution spaces [Covert and Palsson, 2002; Åkesson *et al.*, 2004; Becker and Palsson, 2008; Mo *et al.*, 2009]. The whole feasible solution space can be studied algebraically or statistically by sampling the space [Price *et al.*, 2004a; Palsson, 2006]. Randomized Monte Carlo sampling of the feasible solution space gives unbiased information on the shape and properties of the space where the true metabolic state lies [Price *et al.*, 2004a; Schellenberger and Palsson, 2009]. The null space that contains all the possible flux distribution can be studied algebraically [Palsson, 2006]. Investigation of the feasible solution space yields information on what types of solutions are possible, what parts of the metabolic network participate in the possible metabolic states, are there some limits for production of specific extracellular compounds etc. Obviously the properties of the feasible solution space contain even the properties of the true metabolic state.

Linear optimisation can be utilised to find a point solution, i.e. a single flux distribution. The approach is often called flux balance analysis (FBA) and there the optimisation requires an objective function. It is always a guess what the

organism actually optimises for in the particular conditions and it is generally not simple to set up a biologically meaningful objective function. Biomass production is an obvious choice for objective function for bacteria that grow exponentially. However, higher cells do not usually optimise for growth though uncontrolled growing cancer cells could be an exception. It is generally accepted that organisms have evolved to survive but their survival is not straightforward to define as an objective function. Different types of objective functions have been searched and suggested [Burgard and Maranas, 2003 (ObjFind); Holzhutter, 2004 (flux minimization)] and the optimised flux solutions have been tested against experimental data obtained with ^{13}C-tracer based methods [Schuetz *et al.*, 2007]. ^{13}C-tracer based methods will be presented in the next chapters. In addition to flux determination, FBA approach has been exploited for identification of optimal targets for metabolic engineering [Burgard *et al.*, 2003 (OptKnock); Pharkya *et al.*, 2004 (OptStrain)]. Furthermore, the properties of a metabolic system can be studied by defining different types of objective functions. For example production capabilities can be determined by optimising for the product formation. However, engineered organisms may not initially reach the optimal performance. Thus, FBA will not return flux phenotypes that match the reality of engineered organisms. Successful predictions of flux phenotypes of engineered organisms have been obtained with the minimization of metabolic adjustment (MOMA) to the wild type flux phenotype -principle and solved with quadratic programming [Segré *et al.*, 2002].

2.3.2 ^{13}C-metabolic flux analysis

As discussed above the determination of an objective function for FBA is often extremely difficult. In addition, the constraint-based MFA approaches, like FBA, cannot solve distributions of fluxes to parallel and alternative pathways. However, the parallel pathways usually transfer atoms in distinctive manner before they converge to a common metabolic intermediate. Thereby, utilisation of tracers has emerged. Since metabolism is all about breaking and making carbon-carbon bonds, ^{13}C, is the most common tracer in metabolic studies [Tang *et al.*, 2009]. ^{13}C is a stable carbon isotope whose natural abundance is only 1.1% [Gadian, 1982]. Other tracers are applicable for studies of specific metabolic pathways [Brosnan *et al.*, 2004 (^{15}N tracer application)].

2.3.2.1 ^{13}C-labelling and analytical methods

^{13}C-labelling for MFA is performed by introducing ^{13}C-labelled substrate to a cell culture. When the carbon source gets metabolized the tracer enrichment spreads first to the free intracellular metabolites and during extended periods of growth on ^{13}C-labelled carbon source also into the macromolecules and cell constituents. The spread of the ^{13}C-label is dependent on the metabolic fluxes and the turnover of macromolecules and storage pools. Since the alternative pathways often scramble and cleave the carbon backbones of metabolites in different ways, the ^{13}C-labelling prints information on the relative activities of the pathways into the carbon-carbon connectivities of the metabolites and into the positional fates of tracer atoms. Depending on the difference between the carbon chain modifications in the alternative pathways, different ^{13}C-labelling designs of the carbon source are optimal for resolution of the relative pathway activities [Möllney *et al.*, 1999; Araúzo-Bravo and Shimizu, 2003]. Isotopomer is a definition for isotopic isomers of a compound [Wiechert, 2001] and the isotopomers differ only in position or number of different isotopes in the molecule. Thus, molecules differing in ^{13}C-labelling patterns are isotopomers [Nielsen *et al.*, 2003]. In the following chapters two main ^{13}C-labelling approaches are presented.

Uniform labelling

In uniform ^{13}C-labelling approach the carbon source contains a fraction of molecules that are uniformly ^{13}C-labelled. Thus, they contain ^{13}C atom in all positions. Typically utilised fraction of uniformly ^{13}C-labelled molecules is around 20% [Zamboni *et al.*, 2009]. If the alternative metabolic pathways modify the carbon chain in distinctive ways, the relative activities of the pathways can be resolved with this approach. This approach can even be called bond labelling because the relative activities of the alternative pathways are actually recorded in the common product as cleaved and newly formed carbon-carbon bonds. The uniform ^{13}C-labelling approach was established in 1990's by Szyperski [Szyperski 1995] by introducing biosynthetically directed fractional (BDF) ^{13}C-labelling where a fraction of uniformly labelled carbon source was feed to microbial cells. During the steady-state growth on fractionally ^{13}C-labelled carbon source information on the relative *in vivo* activities of the pathways was recorded and significantly amplified in the ^{13}C-labelling patterns of proteinogenic amino acids. The ^{13}C-labelling patterns of proteinogenic amino acids could conveniently be

detected by two-dimensional nuclear magnetic resonance (NMR) spectroscopic experiments. Szyperski (1995) further introduced probabilistic equations for the relations between the ^{13}C-^{13}C-couplings in proteinogenic amino acids and fragmentomers, fractions of intact carbon fragments. Four fragmentomers that sum up to one of three carbon fragment of a molecule can be deduced. Fragmentomer $f^{(1)}$ represents the fraction of molecules in which the middle carbon atom and the neighboring carbons originate from different carbon source molecules, fragmentomer $f^{(2)}$ represents the fraction of molecules in which the middle carbon atom and one of the two neighboring atoms originate from the same carbon source molecule, and fragmentomer $f^{(3)}$ represents the fraction of molecules in which the middle carbon atom and both the neighboring carbons originate from the same carbon source molecule. Sometimes, if the end carbons of the three carbon fragments are in different chemical environment, even two different fragmentomers $f^{(2)}$ and $f^{(2^*)}$ can be distinguished by NMR spectroscopic methods. Fragmentomers are actually constraints for a full isotopomer distribution of a molecule. Later a GC-MS based detection method was developed to be compatible with BDF labelling [Fischer and Sauer, 2003].

Positional enrichment

The alternative ^{13}C-labelling approach is to introduce positional label(s). The positional label can be introduced in a specific position or positions of the carbon source and usually all the carbon source is equally ^{13}C-labelled. For example, the common carbon source glucose is commercially available in different compositions of ^{13}C and ^{12}C atoms. However, glucose with ^{13}C-atoms somewhere in the middle of the carbon chain is very expensive to purchase. During the growth on positionally ^{13}C-labelled carbon source, specific positions of product molecules become enriched depending on the *in vivo* activities of pathways. In positional fractional ^{13}C enrichments, the ratios of ^{13}C and ^{12}C atoms in the specific carbon positions of the product molecules provide constraints for the full isotopomer distribution of a molecule [Wiechert, 2001].

Nuclear magnetic resonance spectroscopy

Modern nuclear magnetic resonance (NMR) spectroscopy detects signals of spin possessing nuclei in a strong magnetic field after a radio frequency pulse or a sequence of pulses [Friebolin, 1991]. Nuclei that possess a spin different from

zero have different energy states in a magnetic field. Radio frequency pulses induce transitions between the energy states and create detectable macroscopic magnetisation. The magnetisation induces a current to the receiver coil. This signal is recorded and called free induction decay (FID). The FID is then Fourier transformed from time domain to frequency domain to obtain an NMR spectrum. The limited sensitivity of NMR stems from detecting only the small difference between the populations of nuclei on different energy states. The energy difference is dependent on the strength of the static magnetic field. Therefore, strong magnets are utilised. Nevertheless, NMR spectroscopy provides unlimited potential in the variety of methods that can be utilised for analysis of biological samples.

Spin possessing NMR active nuclei of main interest in analysis of biological samples are: ^{1}H, ^{13}C, ^{15}N, and ^{31}P [Gadian, 1982]. All these nuclei have spin quantum numbers of ½ and thus, have two possible energy states in a magnetic field. Proton is the most sensitive nuclei and ^{1}H NMR spectroscopy is an unbiased method because it can detect all proton containing compounds in a sample. On the other hand ^{31}P has a 100% natural abundance and therefore, it can be utilised for example for selective detection of only phosphorus containing compounds in a complex mixture.^{13}C atoms, whose natural abundance is only 1.1%, can be directly detected by NMR but not the more abundant ^{12}C atoms. Thus, ^{13}C is a suitable tracer for NMR spectroscopic studies.

NMR active nuclei give signals in an NMR spectrum on their characteristic chemical shifts [Friebolin, 1991]. The characteristic chemical shift of a nucleus depends on the nature and the chemical environment of the nucleus. Electrons in the chemical environment cause shielding of the magnetic field and thus, the magnetic field experienced by the nucleus is also dependent on its surrounding electrons. In addition, coupling to other NMR active nuclei through bonds gives rise to signal splitting. Therefore, different molecule structures have specific signal fine structures.

Complex sequences of radio frequency pulses can be designed for advanced NMR spectroscopic experiments. Magnetisation can, for example, be transferred from one type of nuclei to other types of nuclei, which targets the analysis to specific structures of interest. Higher dimensional experiments can be performed for one type nuclei (homonuclear) or for different types of nuclei (heteronuclear) [Croasmun and Carlson, 1994]. Higher dimensional experiments provide also more information about the structures of analytes because nuclei that are covalently bound together or close to each other in space can be identified.

NMR spectroscopic analyses of samples from ^{13}C-labelling experiments can provide two different types of constraints to the isotopomer distribution. If ^{13}C are detected, signal fine structures reveal fractions of couplings to adjacent ^{13}C and ^{12}C nuclei, thus, a ^{13}C-labelling status of a three carbon fragment [Szyperski, 1995]. Sometimes even longer couplings can be resolved. If protons are detected, the signal fine structure reveals the fractional enrichment of ^{13}C in the carbon coupled to the detected proton. The fraction of protons coupled to ^{13}C is observed as split satellite signals on both sides of the signal from the ^{12}C coupled protons [Friebolin, 1991].

Mass spectrometry

The advantage of mass spectrometry (MS) compared to NMR is its higher sensitivity. However, there are fundamental differences in the data that is produced by MS and NMR. MS detects molecules that have distinctive masses but it cannot distinguish between molecules having the same number of ^{13}C atoms but in different positions [Zamboni *et al.*, 2009]. Massisomer (or mass isotopomer) is the definition for isomers that differ in mass [Christensen and Nielsen, 1999]. Gas-chromatography-mass spectrometry (GC-MS) has been the most popular of MS techniques for analysis of samples from ^{13}C-labelling experiments [Wittmann, 2007]. Before the GC-MS analysis the metabolites are first derivatized to render the molecules volatile [Tang *et al.*, 2009]. Common derivatizations are silylation, acylation and alkylation [Tang *et al.*, 2009; Wittmann, 2007]. In GC-MS analytes become fragmented and derivatization agent may affect the fragmentation sites. Fragmentation yields more constraints to the full isotopomer distribution [Zamboni *et al.*, 2009]. Metabolites include natural isotopes and derivatization introduces additional atoms to the analytes. Thus, the raw data requires correction to remove them [Christensen and Nielsen, 1999; van Winden *et al.*, 2002].

Previous analyses of ^{13}C-labelling experiments, with NMR spectroscopy or GC-MS, have utilised the detection of ^{13}C-labelling patterns of proteinogenic amino acids that are abundant. GC-MS has somewhat been utilised also in analysis of free amino acids and organic acids [Wittmann *et al.*, 2002]. However, there is a delay before the label reaches proteins that are macromolecules or even some delay before the label enriches in the large free amino acid pools in cells. Therefore, the methods have not been suitable for analysis of phenomena in short time frames. In addition, the long ^{13}C-labelling experiments require lot

of expensive labelled carbon source. The application of liquid chromatography-mass spectrometry (LC-MS) in analysis of ^{13}C-labelling patterns in intracellular metabolic intermediates was demonstrated by van Winden *et al.* (2005) but the direct analysis of ^{13}C-labelling patterns of intracellular metabolites has not become a widespread method because the sensitivity severely limits the analysis as metabolic intermediates are present only in very low amounts. Later Toya *et al.* (2007) suggested CE-TOFMS (capillary electrophoresis time-of-flight mass spectrometry) for detection of ^{13}C-labelling patterns in free intracellular metabolites. CE-TOFMS is fast and the experimental set up is more flexible for analysis of various compounds than LC-MS. Kleijn *et al.* (2007) showed that the data sets from the three measurement techniques: NMR spectroscopy, LC-MS and GC-MS yielded consistent flux results in analysis of combined substrate labelling, 10% [U-^{13}C] and 90% [1-^{13}C] glucose, in glycerol over-producing *S. cerevisiae* strains. Since the flux sensitivities were found to often depend on the analysis method, a combined data set gave the most accurate flux distribution estimate. LC-MS was utilised for detection of ^{13}C in free metabolic intermediates whereas the NMR spectroscopy and GC-MS analyses were performed for ^{13}C-labelling patterns in proteinogenic amino acids and storage carbohydrates.

In order to obtain massisomer data on smaller fragments or even pure positional ^{13}C enrichment data, liquid chromatography-tandem mass spectrometry (LC-MS/MS) was introduced to ^{13}C-labelling analysis task [Iwatani *et al.*, 2007]. In LC-MS/MS the full massisomers are further fragmented and positional enrichments can be inferred from the full fragmentation data. Iwatani *et al.* (2007) applied LC-MS/MS detection to analysis of ^{13}C-labelling patterns of proteinogenic and free amino acids in *E. coli*.

2.3.2.2 ^{13}C-metabolic flux analysis – mathematical and statistical methods

The interpretation of data from ^{13}C-labelling experiments requires mathematical modelling and statistical analysis. Firstly the atom transfers in the metabolic reactions are essential to be modelled for interpretation of data form tracer experiments. Mappings of carbon atoms can be obtained from few sources [Arita, 2003; Kotera *et al.*, 2004, (KEGG rpair); Mu *et al.*, 2007 (carbon fate maps)]. Unfortunately the databases may contain errors or inconsistencies and thus, the mappings for flux analysis models require curation or at least an inspection. If the atom transfers of interest are not found in the databases, one is forced to go into mecha-

nisms of the reactions to resolve them. Furthermore, label scrambling in symmetrically reacting compounds need to be taken into account [Bernhard and Tompa, 1990]. Software such as ReMatch (http://sysdb.cs.helsinki.fi/ReMatch/) [Pitkänen *et al.*, 2008] and OpenFLUX [Quek *et al.*, 2009] aid in setting up and sharing metabolic models that include atom mappings.

The flux estimation methods that exploit ^{13}C-labelling data can be divided into two sub categories: global iterative fitting and local flux ratio analysis (possibly followed by direct flux estimation) methods. In the following chapters the features of both types of computational methods and established protocols and software are presented.

Global iterative fitting

Global iterative fitting requires modelling of label propagation in the metabolic network and set up of balance equations generally for each isotopomer. For a metabolite with n carbons there will be 2^n possible isotopomers. There will be an extremely high number of isotopomer balance equations in the system and many of them are nonlinear. Iterative fitting searches for the best fit between the observed and simulated labelling patterns [Wiechert *et al.*, 2001; Antoniewicz *et al.*, 2006]. Iteration is initiated from a guessed or a random flux distribution and ^{13}C-labelling patterns of metabolites are simulated with the model. The simulated ^{13}C-labelling patterns are compared to the observed ones and the iteration is continued until a minimum of the difference or a difference under a threshold between the simulated and the observed ^{13}C-labelling patterns is reached. The method returns a single flux distribution that gives the best global fit to all the measured data that was utilised as input.

Other methods to model ^{13}C-labelling patterns than isotopomers have been developed since the original task of simulating numerous isotopomers is computationally highly demanding. Transformation of isotopomers into cumomers enabled analytical solution as solving cascades of linear equations [Wiechert *et al.*, 1999]. Cumomers are by definition certains sums of isotopomers and cumomer fractions can incorporate both positional enrichments and isotopomer fractions. Bondomers were introduced for modelling of label propagation in uniform ^{13}C-labelling experiments [van Winden *et al.*, 2002]. Bondomers are isomer entities that differ only in numbers and positions of intact carbon-carbon bonds. Utilisation of bondomers instead of isotopomers or cumomers decreases the number of mass balance equations and similarly as isotopomers can be trans-

formed into cumomers, bondomers can be transformed into cumulative bondomers. Like cumomers cumulative bondomers enable analytical solution of sequence of linear equations. Recently an elementary metabolite units (EMUs) framework was introduced to further reduce the computational time required to simulate isotopic labelling patterns [Antoniewicz *et al.*, 2007]. EMUs are any distinct subsets of metabolites' atoms and the reaction network is decomposed into EMU reactions and a minimum amount of information required for simulations is identified. EMU framework is compatible for simulation of any type of isotopic labelling.

^{13}C-FLUX software was for a long time the only publicly available software framework for ^{13}C-MFA [Wiechert *et al.*, 2001]. It is compatible with all kinds of measurement data and provides also statistical algorithms for analysing the results. EMU framework is utilised in OpenFLUX, which is recent user-friendly software for all the steps of ^{13}C-MFA, from model building to statistical analyses [Quek *et al.*, 2009].

The drawbacks of the iterative fitting methods are that it is difficult to assure that the method reached a global minimum instead of just a local one [Ghosh *et al.*, 2005]. Moreover, if there is not enough data the method returns merely random points from the solution space but still cannot define the feasible solution space.

Local flux ratio analysis

Local flux ratio analysis utilises directly the ^{13}C-labelling data to deduce ratios of converging fluxes in the network. Thus the inaccuracies in the data or in the assumptions or errors in the network model affect the results only locally in contrast to the global methods [Zamboni *et al.*, 2009]. Algebraic equations that relate the ^{13}C-labelling pattern of a junction metabolite to the relative fluxes through the branching pathways are formulated. If it is possible to solve relative fluxes for every pair of alternative pathways in the network model i.e. as many as there are degrees of freedom in the stoichiometric model, then the determined flux ratios as additional constraints render the MFA system solvable. Approach was introduced by Fischer *et al.*, (2004) and has been implemented as software FiatFlux [Zamboni *et al.*, 2005]. The frameworks for flux ratio analysis are described in the following paragraphs.

Metabolic flux ratio analysis

Metabolic flux ratio (METAFoR) analysis was initially developed to rely on uniform ^{13}C-labelling approach by the biosynthetically directed ^{13}C-labelling of the proteinogenic amino acids and following analysis of ^{13}C-labelling patterns by two-dimensional NMR spectroscopy [Szyperski 1995; Szyperski *et al.*, 1999]. Since the carbon backbones of metabolic intermediates of central carbon metabolism are conserved in synthesis of proteinogenic amino acids and the amino acid synthesis pathways were well known for *E. coli*, Szyperski (1995) back propagated the ^{13}C-labelling patterns from the amino acids to metabolites and derived equations for ratios of converging fluxes in central carbon metabolism. Later Maaheimo *et al.* (2001) extended the method and derived flux ratio equations for compartmental metabolism of eukaryotic *S. cerevisiae*. The ^{13}C-labelling patterns of eight metabolic intermediates of central carbon metabolism of *S. cerevisiae* can be determined and utilised as parameters in the flux ratio equations. The equations derived for eukaryotic metabolism have then been utilised in analysis of metabolic states of at least the following other yeasts and a fungus *P. pastoris* [Sola *et al.*, 2004], *P. stipitis* [Fiaux *et al.*, 2003], *P. anomala* [Fredlund *et al.*, 2004] and *T. reesei* (Publication III).

The original analytical method in METAFoR analysis was ^{1}H-^{13}C HSQC (heteronuclear single quantum coherence) NMR spectroscopic experiment [Croasmun and Carlson, 1994] were the proton bound ^{13}C-nuclei and the ^{13}C-labelling status of the adjacent carbon nuclei are detected. In the experiment, signals from ^{13}C-nuclei in proteinogenic amino acids are spread into two-dimensions and found at characteristic chemical shifts in proton and carbon dimensions. According to the ^{13}C-labelling status of the adjacent carbon nuclei, different signal fine structures are formed. Coupling to an adjacent ^{13}C-nucleus splits the signal (Figure 3). All the different signal fine structures are observed on top of each other and their fractional volumes correspond quantitatively to the fractions of different three carbon isotopomers with a central ^{13}C nucleus. The fractions of different three carbon isotopomers with a central ^{13}C nucleus are obtained by iterative fitting of simulated signal fine structures on the whole multiplet signal. FCAL is a software developed for the iterative fitting and following calculation of fragmentomers from the fractions of different signal fine structures with the equations derived by Szyperski (1995) [Szyperski *et al.*, 1999]. The probabilistic equations take into account the fraction of uniformly ^{13}C-labelled substrate, the natural ^{13}C

abundance in the rest of the carbon source and the fraction of biomass synthesized during the ^{13}C-labelled feed [Szyperski *et al.*, 1995].

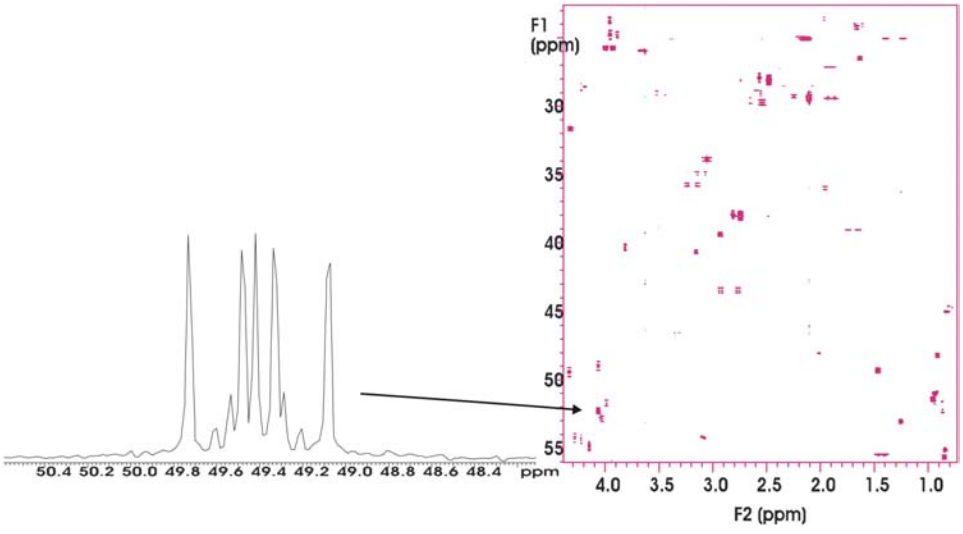

Figure 3. An example of the ^{13}C finestructure of Glu-Cα extracted from a two-dimensional ^{1}H-^{13}C HSQC NMR spectrum.

The sample preparation for METAFoR analysis is simple [Szyperski *et al.*, 1999]. The harvested biomass is just hydrolysed (6 M HCl, +110 °C). Since the information on the *in vivo* fluxes is naturally amplified in biomass, sensitivity does not limit the NMR measurement. For NMR analysis ash is removed from the hydrolysed sample and the solvent is switched to D_2O. The detection of the ^{13}C-labelling patterns of proteinogenic amino acids by 2D NMR spectroscopy can be performed without any separation steps in sample preparation.

Fischer and Sauer (2003) extended the flux ratio analysis to GC-MS as analytical method and initiated the utilisation of mixed ^{13}C-labelling, a combination of positional and fraction uniform ^{13}C-labelling. The introduction of positional label was meaningful in combination of the switch to GC-MS analysis, because even though GC-MS cannot directly quantify positional enrichments, enrichment of label in carbon fragments can be detected. Later the GC-MS analysis has been extended to analysis of ^{13}C-labelling in cell cultures in 1 ml deep-well microtiter plates enabling high throughput metabolic flux profiling [Fischer *et al.*, 2004].

Conventionally the flux ratio equations have been manually derived by experts of the metabolism of the organism under study being able to set up mean-

ingful assumptions of the fluxes for every new organism, metabolic network, ^{13}C-labelling of the carbon source and an analytical platform. Only recently a framework for systematic derivation of the equations from a given metabolic network model was introduced (Publication IV).

Local flux ratios as additional constraints in metabolic flux analysis

Local flux ratios determined from ^{13}C-labelling experiments are experimental information that can be utilised as additional constraints in a conventional MFA system [Fischer *et al.*, 2004]. If every branching point in the network model can be constrained, the system renders solvable. Fischer *et al.* (2004) applied the approach to metabolism of *E. coli*. Constraint equations were set up and the net fluxes in central carbon metabolism were solved by constrained nonlinear optimisation with Matlab function *fmincon*. Later Fredlund *et al.* (2004) set up the constraint equations for compartmentalised eukaryotic network model of *Pichia anomala*. Both studies utilised GC-MS determined ^{13}C-labelling patterns for flux ratio determination. Zamboni *et al.* (2005) implemented the approach as software FIATFLUX for net flux determination in three organisms: *E. coli, Bacillus subtilis, S. cerevisiae*. FIATFLUX contains two Matlab based modules. The first one is for determining local flux ratios of GC-MS massisomer data from ^{13}C-experiments and the second module is for estimating the net fluxes utilising the local flux ratios as additional constraints. The network models for determining the local flux ratios are fixed but the stoichiometric models for net flux determination are open for users. Constraints from ^{13}C-labelling experiments often enable leaving out the cofactors form metabolite balancing. Cofactor balancing is highly error prone since the cofactor specificities vary in isoenzymes and they are not precisely known. In Publication I net fluxes in central carbon metabolism of *S. cerevisiae* were determined under different oxygenation conditions by performing ^{13}C-labelling experiments, utilising NMR spectroscopy based METAFoR analysis instead of GC-MS analysis for determining the local flux ratios and then solving the net fluxes by optimisation having the flux ratios as additional constraints.

2.3.2.3 ^{13}C-metabolic flux analysis in large scale networks

In prolonged ^{13}C-labelling experiments all cell components and metabolites become ^{13}C-labelled but analysis of ^{13}C-labelling patterns of only limited number of metabolites is feasible with current MS and NMR techniques. However, the

measurements with both methods can be directed to specific compounds or methods can be adjusted for analysis of desired compounds. Thus, computational ^{13}C-experiment design has been investigated [Möllney *et al.*, 1999]. In addition to optimal ^{13}C-labelling designs for flux resolution [Möllney *et al.*, 1999; Araúzo-Bravo and Shimizu, 2003], sets of most informative compounds for ^{13}C-MFA can be computationally determined to target the analysis of ^{13}C-labelled samples [Rantanen *et al.*, 2006].

^{13}C-MFA becomes unfeasible in large-scale networks because of limited ^{13}C-labelling data and because of the size of the equation system. For iterative methods the number of equations grows fast as the number of additional isotopomer balances explodes when the network size increases. ^{13}C-MFA with local flux ratios as additional constraints would be computationally feasible in larger networks if it was possible to determine sufficiently many local flux ratios. Thus, the ^{13}C-labelling data limits that approach.

However, few flux analysis studies in large-scale networks have been carried out. Blank *et al.* (2005) determined the net fluxes in the central carbon metabolism of *S. cerevisiae* with local flux ratios as additional constraints in MFA and estimated the fluxes in large-scale network by minimisation of fluxes outside the central carbon metabolism. Quite recently Suthers *et al.* (2007) applied iterative flux determination approach to a large scale network of *E. coli*. Even though they included both cofactor balancing and ^{13}C-labelling data, multiple local optima that were statistically indistinguishable were identified. The ^{13}C-labelling data was insufficient for reliable flux determination in the large-scale model.

3. Aims of the research

The research included in the thesis has concentrated on studying the metabolism of mainly three organisms: yeasts *S. cerevisiae* and *P. pastoris* and a fungus *T. reesei* all of which are important production organisms in biotechnology. The focus has been on the process streams of the cell factories, the metabolic fluxes, under conditions of interest for development of production processes. The quantitative studies of the distribution of intracellular fluxes of different organisms and under different growth conditions have required development of modelling of metabolism for the analysis of fluxes (Publications I, II, III). In Publication IV and somewhat also in Publication I, computational methods for quantitative analysis of the intracellular fluxes were developed. In Publication V a novel step in the fungal metabolic pathway of catabolism of plant material compound D-galacturonic acid was identified and thus a previously unknown reaction node and interactions to reactant metabolites were set in the fungal metabolic network. The analytical tool utilised in the detection of ^{13}C-labelling in the analyses of metabolic fluxes, namely the NMR spectroscopy, was applied to verify the product of the novel enzyme.

3.1 Oxygen dependence of fluxes and underlying regulation in *S. cerevisiae*

Understanding the regulation of redox homeostasis of the organism is important for any metabolic engineering project because redox homeostasis binds together functions of even distant pathways. Redox homeostasis of an organism is naturally affected by the oxygenation of the culture that is a major parameter in industrial bioprocesses. Oxygenation is also one of the factors that most contribute to the cost of a bioprocess. The important bioproduction and model organism *S. cerevisiae* is known to exhibit various states of energy metabolism depending

on the prevailing growth conditions. However, the effect of different low oxygen conditions on the metabolic state of *S. cerevisiae*, thus the distribution of the metabolic fluxes, has not been thoroughly studied. Thus, the aim of the Publication I was to quantify the dependence of the intracellular flux distributions of *S. cerevisiae* on oxygen provision and to study the maintenance of redox homeostasis in the different oxygenation conditions. It was essential to quantify the intracellular net fluxes without including the redox cofactors in the metabolite mass balances, because the conditions were expected to strongly affect the redox balancing. Therefore, in Publication I ^{13}C-MFA was utilised to determine the net fluxes.

3.2 Two carbon source case of methanol and glycerol utilisation by *P. pastoris*

Methylotrophic yeast *P. pastoris* is a host organism for industrial production of heterologous proteins. Strong inducible promotors of the genes of methanol utilisation pathway in the peroxisomes are utilised to induce the expression of recombinant proteins. Thus, metabolism during the co-utilisation of inducer methanol and a carbon source is of interest from process development point of view. The METAFoR analysis of *P. pastoris* [Sola *et al.*, 2004] was extended to a two-carbon source co-utilization. The eukaryotic model for METAFoR analysis [Maaheimo *et al.*, 2001] was likewise extended with methanol utilisation pathway. The ^{13}C-labelling with the uniform labelling approach was performed in continuous cultures of *P. pastoris* growing on different methanol/glycerol mixtures and at two growth rates to probe the intracellular metabolic state, the ratios of intracellular metabolic fluxes, in different possible process conditions. The aim was to provide valuable information for process optimisation of recombinant protein production with *P. pastoris*.

3.3 Path identification and the effect of carbon catabolite repression on metabolic fluxes in *T. reesei*

The efficient protein expression machinery of *T. reesei* has been widely investigated but even despite the wide industrial importance the metabolism of *T. reesei* has not been largely studied and its potential is still not known. The genome of *T. reesei* has recently been published [Martinez *et al.*, 2008] but the fungus still lacks a genome-wide metabolic reconstruction.

The transcriptional response to preferred and repressive carbon source glucose and to a neutral carbon source has been studied by Chambergo *et al.* (2002). They found that excess glucose does not repress the respiratory pathway genes to the similar extent as in *S. cerevisiae* that is adapted to fast utilization of sugars by fermentation in nutrient rich environments. However, the intracellular *in vivo* fluxes of *T. reesei* were not previously studied and the effect of the different conditions of carbon catabolite repression on the intracellular fluxes has not been known.

Because of the lack of a metabolic reconstruction of *T. reesei* the biosynthetic pathways of proteinogenic amino acids in *T. reesei* were reconstructed with a recent computational pathway analysis method ReTrace [Pitkänen *et al.*, 2009] in Publication III. The reconstruction of the biosynthetic pathways of proteinogenic amino acids was essential for the application of the METAFoR analysis to quantitatively probe the intracellular flux distributions in *T. reesei*. The ^{13}C-labelling of *T. reesei* for METAFoR analysis was performed for cells growing on preferred and repressive carbon source glucose and on sorbitol. The effect of induction of cellulase gene expression on the ratios of intracellular fluxes was also studied by ^{13}C-labelling of a *T. reesei* culture growing on sorbitol induced with a small addition of inducer sophorose.

3.4 Framework for analytical determination of flux ratios

The previously established computational methods for quantitative analysis of the metabolic fluxes from ^{13}C-isotopomer measurement data relied either on manual derivation of analytic equations constraining the fluxes or on numerical solution of a highly nonlinear system of isotopomer balance equations. In the first approach, analytic equations were to be tediously derived for each organism, particular growth conditions and substrate or labelling pattern, by a domain expert while in the second approach, the global nature of an optimum solution is difficult to prove and comprehensive measurements of external fluxes to augment the ^{13}C-isotopomer data were typically required. A framework for an automatic and systematic derivation of equation systems constraining the fluxes from the model of the metabolism of an organism was developed. The framework was designed to be general for all metabolic network topologies, ^{13}C-isotopomer measurement techniques, carbon sources, and carbon source ^{13}C labelling patterns.

3.5 NMR spectroscopy as a tool in pathway identification

Fungal metabolism comprises a lot of hitherto unknown potential for biotechnology. For example a fungal pathway for D-galacturonic acid catabolism has been only partly known. D-galacturonic acid is a major component of pectin that is abundant in plant material. The first step in the fungal pathway of D-galacturonic acid catabolism was previously identified and it is an NADPH-specific D-galacturonic acid reductase generating L-galactonate [Kuorelahti *et al.*, 2005]. The next reaction in the pathway, a novel enzyme that converts L-galactonate to L-threo-3-deoxy-hexulosonate was then identified in *T. reesei*. The active enzyme was produced in the heterologous host *S. cerevisiae* and characterized. The reaction product of the enzyme L-galactonate dehydratase was analysed and identified by NMR spectroscopy. 1D and 2D NMR spectroscopic experiments were utilised for the identification of the reaction product directly in the reaction mixture.

4. Research methods

4.1 Strains

In Publication I *S. cerevisiae* CEN.PK113-1A (MATα, URA3, HIS3, LEU2, TRP1, MAL2-8c, SUC2) strain was employed. The strain was kindly provided by Dr. P. Kötter (Institut für Mikrobiologie, J.W. Goethe Universität Frankfurt, Germany) [de Jong-Gubbels *et al.*, 1998] and prior to the experiments stored in glycerol (30% v/v) at -80 °C [Wiebe *et al.*, 2008]. In Publication II a prototrophic *P. pastoris* strain expressing a heterologous protein, a *Rhizopus oryzae* lipase, under the transcriptional control of the *aox-1* promoter was employed. *P. pastoris* X-33/pPICZαA-ROL [Minning *et al.*, 2001] is the wild-type phenotype X-33 strain (Invitrogen) with the pPICZαA-derived expression vector (Invitrogen) containing the *ROL* gene, pPICZαA-ROL, integrated in its *aox-1* locus. In Publication III *T. reesei* strains QM6a (wild type) [Mandels and Reese, 1957] and QM6a with deleted *cre1* gene (unpublished) were employed. In Publication V the *S. cerevisiae* strain CEN.PK2-1D (VW-1B) was employed as the host for the heterologous expression of a *T. reesei* enzyme and was the source of the extract in the NMR spectroscopic analyses. *T. reesei* strains Rut C-30 or QM6a were employed otherwise.

4.2 Cultivations

In Publications I, II and III the organisms were cultivated either in continuous cultures in fermentors or in batch cultures in flasks. Continuous cultivations provide highly controlled culture conditions were a single parameter can be varied while everything else is kept constant. The growth rate of an organism is set by the rate of the feed and by the flow out of the reactor. Continuous cultivation operates in steady state continuous mode when all the variables have time inde-

pendent values. The steady state continuous mode is usually eventually obtained when after the initiation of the culture in a batch mode all the feed variables are kept constant. The reactor is assumed to be an ideal bioreactor where the properties of the effluent are identical to the properties of the culture in any point of the reactor. The assumption is reasonable when the mixing of the reactor is efficient. Because of the strict control, continuous cultivation provides possibilities for sampling during steady state conditions and highly reproducible cultures. The continuous cultivations utilized are described in more detail in Publications I and II and the cultivation set up in Publication I also in Wiebe *et al.* (2008).

Batch cultures in flasks are much simpler to prepare than continuous cultures. The filamentous growth of *T. reesei* complicates the bioreactor cultivations and thus in Publication II *T. reesei* was cultivated in flasks. After the initial lag-phase, during the early exponential phase in a flask culture, the growth conditions are still almost unchanged from the initial culture conditions and the organism is growing on its maximal growth rate. After passing the early exponential phase the growth conditions are not precisely known anymore. The *T. reesei* batch cultures in flasks are described in detail in Publication III.

All the cultivations for METAFoR analysis in Publications I, II, III were performed on minimal medium without amino acids. The media of the main cultures were as follows. Information on the media for inoculates etc can be found in Publications II and in Wiebe *et al.* (2008) for cultures in Publication I. In Publication I yeast was grown on defined minimal medium [Verduyn *et al.*, 1992], with 10 g glucose l^{-1} as carbon source, and supplemented with 10 mg ergosterol l^{-1} and 420 mg Tween 80 l^{-1} (a source of oleic acid). In Publication II *P. pastoris* was grown on minimal medium containing (per $1x10^{-3}$ m^3 of deionized water): Yeast Nitrogen Base (YNB), $0.17x10^{-3}$ kg; $(NH_4)_2SO_4$, $5x10^{-3}$ kg; glycerol and methanol (different ratios on w/w basis), $10x10^{-3}$ kg (total). In Publication III *T. reesei* was grown on minimal medium: $(NH_4)_2SO_4$ 7.6 g/l, KH_2PO_4 15.0 g/l, 2.4 mM $MgSO_4$, 4.1 mM $CaCl_2$, $CoCl_2$ 3.7 mg/l, $FeSO_4 \cdot 7H_2O$ 5 mg/l, $ZnSO_4 \cdot 7H_2O$ 1.4 mg/l, $MnSO_4 \cdot 7H_2O$ 1.6 mg/l, pH adjusted to 4.8 with KOH, supplemented with 2% (w/v) carbon source glucose or sorbitol.

4.3 Biosynthetically directed fractional ^{13}C-labelling

Biosynthetically directed fractional (BDF) ^{13}C-labelling was performed for *S. cerevisiae*, *P. pastoris* and *T. reesei* in studies in Publications I, II and III, respectively. *P. pastoris* and *S. cerevisiae* ^{13}C-labellings were performed in

chemostats while *T. reesei* was ^{13}C-labelled in batch cultures in flasks. After reaching a metabolic steady state in *S. cerevisiae* glucose-limited chemostats, as determined by constant physiological parameters including biomass production, carbon dioxide evolution and oxygen uptake rates (CER and OUR), alkali utilisation, and subsequently confirmed by the observation of constant extracellular and intracellular metabolites and gene transcription, 10 % of the carbon source in the medium was replaced with [U-^{13}C]glucose (Publication I). 10% [U-^{13}C]glucose was fed for 1.5 volume changes. *P. pastoris* was cultivated in chemostat on different glycerol/methanol mixtures until a metabolic steady state was reached as indicated by a constant cell density and constant oxygen and CO_2 concentrations in the bioreactor exhaust gas (Publication II). Then the culture was fed with medium containing about 10% (w/w) uniformly ^{13}C-labelled and 90% unlabelled amounts of each substrate simultaneously for one volume change. *T. reesei* ^{13}C-labelling was performed with two different set ups (Publication III). In the first set up ^{13}C-labelled carbon source was provided directly in the medium that was inoculated: in 2% (w/v) glucose minimal medium containing 10% (w/w) [U-^{13}C]glucose and in 2% (w/v) sorbitol minimal medium containing 10% (w/w) [U-^{13}C]sorbitol. In the second set up in exponential growth phase the six flask cultures were combined and then the culture broth was divided into six flasks. The final concentration of 1mM sophorose was introduced into three of the six replicate 2 l flasks to induce cellulase gene expression in *T. reesei*. An identical volume of water was added to the three control cultures. Three hours after the induction, when cellulase gene expression was expected to be at a moderate level [Ilmén *et al.*, 1997], 0.4 g of [U-^{13}C]sorbitol was added to all six cultures to initiate BDF ^{13}C-labelling. The addition of 0.4 g of [U-^{13}C]sorbitol at this time was estimated to result in a [U-^{13}C]sorbitol fraction of about 10% of the total sorbitol in the culture medium.

During steady state growth in chemostats and during quasi-steady state growth in the exponential growth phase in batch cultures the intracellular metabolic fluxes remain constant and determine the labelling patterns of carbon backbones of proteinogenic amino acids formed in biosynthesis. In the *T. reesei* cultures that were inoculated in fractionally ^{13}C-labelled medium the fraction of the initial biomass could be neglected and all the biomass is formed during ^{13}C-labelling (Publication III). Also when the ^{13}C-label was introduced after the induction of cellulase gene expression the initial biomass could be neglected. However, in chemostat cultures the fraction of biomass formed during the ^{13}C-labelling depends on the growth rate (equals dilution rate D) and the duration of the

^{13}C-labelled feed and can be estimated from the first order wash-out kinetics (Publications I and II). The *S. cerevisiae* chemostat cultures at D = 0.1 h^{-1} were fed with fractionally ^{13}C-labelled carbon source for 1.5 residence times which equals 15 h with the particular growth rate (Publication V). The *P. pastoris* chemostat cultures at D = 1.39 x 10^5 s^{-1} and 4.44 x 10^5 s^{-1} were fed with fractionally ^{13}C-labelled carbon source for 1.0 residence time (Publication II).

4.4 Sampling

The biomass samples of *S. cerevisiae* and *P. pastoris* were harvested by centrifugation (Publications I and II, respectively) and *T. reesei* mycelium by filtration (Publication III). The cell pellets and the filtrated mycelium were suspended into 10 ml of 6 M HCl and the biomass was hydrolysed in sealed glass tubes at +110 ºC for 22 h. The suspensions were dried and dissolved in H_2O for filtration through 0.2 μm filters. The filtrates were vacuum-dried and dissolved in D_2O for NMR experiments. The pH of the samples was below 1 due to residual HCl.

4.5 NMR spectroscopy

In Publication V the reaction mixture of *S. cerevisiae* extract of the strain expressing the lgd1 from *T. reesei* and 110 mM L-galactonate was analysed by NMR after different time intervals. The reaction product was identified by comparing the NMR spectrum of the reaction mixture with the NMR spectrum of pure L-galactonate. The NMR experiments were carried out at +23 °C on a Varian Inova spectrometer operating on a proton frequency of 500 MHz. The spectral widths of the 1D ^{1}H and ^{13}C spectra were 5000 Hz and 30 675 Hz, respectively. In two-dimensional homonuclear correlation spectroscopy (COSY) and total correlation spectroscopy (TOCSY) experiments [Croasmun and Carlson, 1994], the spectral widths were 3400 Hz. The spinlock time for magnetization transfer along coupled nuclei in the TOCSY was 80 ms. In two-dimensional heteronuclear ^{1}H-^{13}C HSQC experiment the spectral widths in ^{1}H and ^{13}C dimensions were 1654 Hz and 10 000 Hz, respectively.

For METAFoR analysis in Publications I, II and III the ^{1}H-^{13}C HSQC NMR spectra [Croasmun and Carlson, 1994] were acquired at +40 ºC on a Varian Inova spectrometer operating at on a proton resonance frequency of 600 MHz essentially as described [Szyperski, 1995]. For each sample two spectra were acquired focusing on the aliphatic and aromatic regions. The spectral widths in

the aliphatic spectra were 6000 Hz and 5100 Hz in the ^{1}H and ^{13}C dimensions, respectively. The narrow spectral width in the ^{13}C dimension leads to back-folding of part of the signals to the empty regions of the spectrum. The spectral widths for the aromatic spectra were 6000 Hz and 2815 Hz in the ^{1}H and ^{13}C dimensions, respectively. The spectra were processed using the standard Varian spectrometer software VNMR (version 6.1, C).

4.6 Metabolic flux ratio analysis

The software FCAL (R.W. Glaser; FCAL 2.3.1) [Szyperski *et al*., 1999] was used for the integration of ^{13}C-scalar fine structures of proteinogenic amino acid carbon signals in the ^{1}H-^{13}C HSQC NMR spectra and the calculation of relative abundances of intact carbon fragments originating from a single source molecule of glucose as in Szyperski (1995).

The nomenclature utilised for the intact carbon fragments, fragmentomers, was initially described by Szyperski (1995) and has briefly been explained in Introduction. Since the carbon backbones of eight metabolic intermediates are conserved in amino acid synthesis, fragmentomer information obtained from proteinogenic amino acids can be traced back to the intermediates of central carbon metabolism [Szyperski 1995; Maaheimo *et al.*, 2001]. Mass balance equations of specific carbon fragments of the metabolic intermediates can be formulated from the propagated fragmentomer information to solve ratios of fluxes in junctions of central carbon metabolism. In Publication I the metabolic flux ratio (METAFoR) analysis relied on the compartmentalized metabolic model of *S. cerevisiae* central carbon metabolism and some of the flux ratios were calculated as formulated by Maaheimo and co-workers (2001). However, some flux ratio calculations were redefined as follows. The nomenclature of the metabolites with differentially conserved C-C connectivities that have been back propagated from the ^{13}C-labelling patterns of amino acids is <metabolite_abbreviation_> and then the following characters denote the status of the bonds in the carbon chain of the metabolite: 1 stands for an intact bond, 0 for a cleaved bond and x for either of the two. The corresponding amino acids fragmentomers are named as was explained above.

The fraction of Pep originating from phosphoenolpyruvate carboxykinase activity, denoted by X_{PEPck}, was calculated from the ratio of the fraction of Pep molecules containing an intact C1-C2 fragment and a cleaved bond between C2 and C3 (Pep_10) and the fraction of Oaa_{cyt} molecules containing the equivalent

fragments (Oaa_{cyt}_10x) (Equation 1). These fragments cannot originate from glycolysis or from the PPP [Maaheimo *et al.*, 2001]. Phe-Cα, Tyr-Cα and Asp-Cα, Thr-Cα can be traced back to the C2 of Pep and Oaa_{cyt}, respectively [Maaheimo *et al.*, 2001] (Equation 6).

$$X_{PEPck} = Pep\_10/Oaq_{cyt}\_10x = \left[f^{(2*)}\right]\{Phe,Tyr-C\alpha\}/\left[f^{(2*)}\right]\{Asp,Thr-C\alpha\} \quad (6)$$

The Oaa_{mit} molecules originating from Oga through the TCA cycle possess cleaved C2-C3 bonds. The fraction of Oaa_{mit} originating from transport over the mitochondrial membrane from Oaa_{cyt} was solved from a mass balance of intact C2-C3 fragments in Oaa_{mit}. The conserved connectivity of the C2-C3 fragment in Oaa_{mit} could be propagated back from Glu-Cα and Pro-Cα carbons that represent the C2 carbon in Oga, since the C2-C3 fragment of Oaa_{mit} is conserved in the TCA cycle as the C2-C3 fragment of Oga. The fraction of Oaa_{mit} from Oaa_{cyt}, denoted by $X_{Oaa-transport}$, was calculated as a ratio of intact C2-C3 fragments in Oga and Oaa_{cyt} (Equation 7).

$$\begin{aligned} X_{Oaa-transport} &= Oga\_x1xx/Oaa_{cyt}\_x1x \\ &= \left(\left[f^{(2)}+f^{(3)}\right]\{Glu,\mathrm{Pro}-C\alpha\}\right)/\left(\left[f^{(2)}+f^{(3)}\right]\{Asp,Thr-C\alpha,Asp-C\beta\}\right) \end{aligned} \quad (7)$$

The fraction of Oaa_{cyt} originating from Pyr_{cyt}, denoted by $X_{Oaa_{cyt}\_from\_Pyr_{cyt}}$, was solved from the mass balance of intact C2-C3 fragments (Equation 8). Since the flux from Pep to Pyr_{cyt} through phosphoenolpyruvate kinase and further through pyruvate carboxylase to Oaa_{cyt} could be assumed to be irreversible, the C2-C3 fragments of Pep were used in the mass balance equations. The conserved connectivity of the C2-C3 fragment in Pyr_{cyt} could be observed from Phe-Cα and Tyr-Cα that represent the C2 carbon of Pep (Equation 8).

$$\begin{aligned} X_{Oaa_{cyt}\_from\_Pyr_{cyt}} &= \left(Oaq_{cyt}\_x1x - Oga\_x1xx\right)/\left(Pep\_x1 - Oga\_x1xx\right) \\ &= \frac{\left[f^{(2)}+f^{(3)}\right]\{Asp,Thr-C\alpha,Asp-C\beta\}-\left[f^{(2)}+f^{(3)}\right]\{Glu,\mathrm{Pro}-C\alpha\}}{\left[f^{(2)}+f^{(3)}\right]\{Phe,Tyr-C\alpha,C\beta\}-\left[f^{(2)}+f^{(3)}\right]\{Glu,\mathrm{Pro}-C\alpha\}} \end{aligned} \quad (8)$$

The upper and lower bounds for Pyr_{mit} originating from the malic enzyme reaction, denoted by X_{MAE_ub} and X_{MAE_lb} respectively, were calculated from a mass balance of intact C2-C3 fragments of Pyr_{mit} (Equations 9 and 10). The upper and lower bounds were obtained from the assumption that the substrate fragment for malic enzyme has an equally conserved connectivity as Oga and

Oaa_{mit}. The intact fragments in Oaa_{mit} were obtained from the intact fragments in Oga since the C2-C3-C4 fragment of Oaa_{mit} is conserved in the TCA cycle in synthesis of Oga. The intact fragments in biosynthetic precursor Oga were deduced from the *f*-values of Glu and Pro carbons (Equations 9 and 10).

$$X_{MAE\_ub} = (Pep\_x1 - Pyr_{mit}\_x1)/(Pep\_x1 - Oga\_x1xx) = \frac{\left[f^{(2)} + f^{(3)}\right]\{Phe, Tyr - C\alpha, C\beta\} - \left[f^{(2)} + f^{(3)}\right]\{Ala - C\alpha, C\beta\}}{\left[f^{(2)} + f^{(3)}\right]\{Phe, Tyr - C\alpha, C\beta\} - \left[f^{(2)} + f^{(3)}\right]\{Glu, \mathrm{Pr}o - C\alpha\}} \quad (9)$$

$$X_{MAE\_lb} = (Pep\_x1 - Pyr_{mit}\_x1)/Pep\_x1 = \frac{\left[f^{(2)} + f^{(3)}\right]\{Phe, Tyr - C\alpha, C\beta\} - \left[f^{(2)} + f^{(3)}\right]\{Ala - C\alpha, C\beta\}}{\left[f^{(2)} + f^{(3)}\right]\{Phe, Tyr - C\alpha, C\beta\}} \quad (10)$$

4.7 Metabolic modelling for ^{13}C-metabolic flux analysis

In Publication I metabolic flux analysis (MFA) was used to determine intracellular net fluxes of *S. cerevisiae* under different conditions of oxygen provision (20.9%, 2.8%, 1.0%, 0.5% and 0.0% O_2 in the chemostat inlet gas), with METAFoR analysis providing additional experimental constraints to render the MFA system solvable [Fischer *et al*., 2004]. A stoichiometric model of central carbon metabolism of *S. cerevisiae* was formulated. The system boundary was set around the central carbon metabolism and the model thus included the glycolytic and the pentose phosphate pathways, the TCA cycle and the fermentative pathways, production of glycerol and anabolic fluxes from metabolic intermediates to biosynthesis. The glyoxylate cycle was omitted from the model since the METAFoR analysis data showed that the pathway was inactive (Publication I). Separate pools of Pyr, AcCoA and Oaa in the two cellular compartments, cytoplasm and mitochondria, were included in the model. Thus, they were modelled as two distinct metabolites. Mal was lumped in the same pool with Oaa_{mit}. Also the pentose phosphates formed a single pool and the triose phosphates were combined in the pools of G3P and Pep. DHAP, the precursor for glycerol synthesis, was also combined with the G3P pool. Lumping of the metabolite pools in the model is reasonable when it is meaningful to assume fast exchange between the metabolites, faster than between the metabolites and other compounds. TCA cycle metabolites were represented by the pools of citrate, Oga and Oaa_{mit}.

Scrambling of ^{13}C-labels in the symmetric molecules succinate and fumarate was taken into account [Bernhard and Tompa, 1990].

However, despite the symmetry, partial channelling of succinate and fumarate has also been seen [Sumegi *et al.*, 1993]. The transport of Pyr and Oaa across the mitochondrial membrane were included in the model but the transport of AcCoA, the final step of the cytosolic pyruvate dehydrogenase (PDH) bypass, was omitted since exogenous carnitine would be required for the carnitine shuttle to be active [Lange, 2002; Swiegers *et al.*, 2001; van Roermund *et al.*, 1999], and carnitine was not provided in the medium. In addition, carnitine acetyltransferase activity has not been detected in *S. cerevisiae* grown in anaerobic chemostats at 0.1 h^{-1} [Nissen *et al*., 1997]. However, contradictory observation has been made by Frick and Wittmann (2005) in *S. cerevisiae* strain ATCC 32167. Pyruvate by-pass via transport of AcCoA into mitocnodria was observed to be active during both respirative and fermentative growth under different dilution rates between 0.10 h^{-1} and 0.45 h^{-1}. Since acetaldehyde can freely diffuse across the mitochondrial membrane and acetaldehyde dehydrogenase (EC 1.2.1.3) and AcCoA synthetase (EC 6.2.1.1) enzymes have both been isolated in the mitochodrial proteome [Sickmann *et al*., 2003], PDH bypass could also be partially located in mitochondria and contribute directly to the formation of $AcCoA_{mit}$. In absence of fluxes inducing significantly dissimilar labelling patterns to the C2-C3 fragments of Pyr_{cyt} and Pyr_{mit} i.e. in conditions of low malic enzyme fluxes as observed in this study, ^{13}C-labelling cannot solely reveal the possible contribution of PDH bypass pathways to the carbon flux to mitochondria. However, in the cultivations performed, the expression of *ACS1* encoding the mitochondrial AcCoA synthetase, essential for the contribution of mitochondrial PDH bypass to the formation of $AcCoA_{mit}$, was negligible and the expression of *ACS2* encoding the cytosolic isoenzyme was substantially higher [Wiebe *et al.*, 2008]. Therefore, the mitochondrial PDH bypass was not included in the model.

A model of central carbon metabolism of *S. cerevisiae* with the same extent as above was formulated for development of the systematic and analytic framework for determination of flux ratios in Publication IV. In the model, some simplifications common to ^{13}C-MFA were applied by pooling metabolites whose isotopomer pools can be assumed to be fully mixed (cf. [Kleijn *et al.*, 2007]). Pooling of metabolites was carried for the pentose-phosphates in PPP, phopshotrioses between G3P and Pep in glycolysis, and Oaa and Mal in the TCA cycle. In these cases, pooling was justified by the existence of fast equilibrating, bidirectional reactions between the listed intermediates and the empirical evidence that their

isotopic labelling is not distinguishable with the current analytical tools. Cofactor metabolites were excluded from the model as cofactor specificities and activities are not accurately known for many reactions. The bulk of the carbon mappings of reactions in the metabolic network were provided by ARM project [Arita, 2003]. Carbon mappings from amino acids to their precursors conformed to [Szyperski, 1995] and [Maaheimo *et al.*, 2001]. For empirical verification of the framework it was tested by estimation of flux ratios for junction metabolites in the metabolic network of *S. cerevisiae* from the artificial data generated by the 13C-FLUX software [Wiechert *et al.*, 2001].

In Publication I the metabolic fluxes were modelled as net fluxes so that a net flux in the forward direction was assigned with a positive value and a net flux in the reverse direction was assigned with a negative value. As an exception, the transport of Oaa across the mitochondrial membrane was modelled as two one-directional transport reactions by not allowing negative net fluxes. In *S. cerevisiae* the transport of Oaa across the mitochondrial membrane can occur via mitochondrial Oaa transporter OAC1 facilitated transport [Palmieri *et al.*, 1999].

The stoichiometric model for experiments under 20.9%, 2.8% and 1.0% oxygen conditions consisted of 38 reactions coupling 34 metabolites including duplicated extracellular metabolites and uptake and production fluxes (Publication I Figure 4). The 14 fluxes across the system boundary included glucose uptake and excretion fluxes of ethanol, acetate and glycerol and the fluxes of the metabolic precursors to macromolecule synthesis for biomass production. The METAFoR analysis results were used to identify inactive reactions, to constrain the stoichiometric models for the experiments with 0.5% and 0.0% oxygen by omitting inactive fluxes to avoid numerical problems in optimization. The stoichiometric model for experiments under 0.5% oxygen consisted of 37 reactions, coupling 34 metabolites and excluding the malic enzyme activity from the first model of the network of active reactions. The compartmentalization of central carbon metabolism in anaerobic conditions is evident from the vital anabolic role of mitochondria in the absence of oxygen [Visser *et al.*, 1994]. However, in completely anaerobic conditions only the net transport of Oaa across the mitochondrial membrane is resolvable and the activities of PEPck and malic enzyme reactions cannot be quantified. Since, according to the METAFoR analysis, the PEPck reaction showed only slight activity in the other conditions studied and its activity decreased as the oxygen provided was reduced, it was omitted from the anaerobic stoichiometric model. *MAE1* has been shown to be induced in anaerobic conditions and its possible role in provision of NADPH in mitochondria in

anaerobic conditions has been discussed [Boles *et al.*, 1998]. However, the malic enzyme reaction also showed only slight activity in all the conditions where quantification was possible and had its lowest activity in 0.5% oxygen. Thus, the malic enzyme reaction was omitted from the anaerobic model. Under anaerobic conditions the stoichiometric model of the active pathways consisted of 34 reactions and 34 metabolites.

After including the measured uptake and excretion rates and the rates of metabolic precursor depletion to biomass synthesis, as determined from the composition of *S. cerevisiae* biomass previously reported [Gombert *et al.*, 2001], in the models, the linear equation systems remained underdetermined. The composition of *S. cerevisiae* biomass was assumed to be the same in all the conditions studied, since the biomass composition in the two extreme conditions, i.e. in fully aerobic and in anaerobic conditions, has been experimentally shown to be essentially the same [Gombert *et al.*, 2001; Nissen *et al.*, 1997]. Solvable systems were obtained by further constraining the MFA systems with one to six linearly independent constraints, depending on the structure of the network of active reactions from the METAFoR analysis as described by Fischer and coworkers (2004) for MS ^{13}C-labelling data. Using the constraints from the METAFoR analysis, it was not necessary to include redox cofactor mass balances in the mass balance constraints in ^{13}C MFA. Cofactor mass balances are sources of errors since the correct balancing requires detailed knowledge of the relative activities of different isoenzymes and the enzymes' redox cofactor specificities on a cell wide scale under the studied conditions. Under the conditions of different oxygen provisions, the external conditions posed different challenges on the redox homeostasis systems of the cells and their effect are not known. The mass balances of the metabolites were formulated as a linear equation system as described in [Fischer *et al.*, 2004] (Equation 11):

$$\mathbf{N}_i\mathbf{v} - \mathbf{b} = \mathbf{R}_m \tag{11}$$

where $\mathbf{N}_i$ is the stoichiometric matrix of the active network i determined from the METAFoR analysis fragmentomer data, $\mathbf{v}$ is the flux distribution vector, $\mathbf{b}$ is the vector of the measured extracellular fluxes and $\mathbf{R}_m$ is the vector of the residuals of metabolite mass balances.

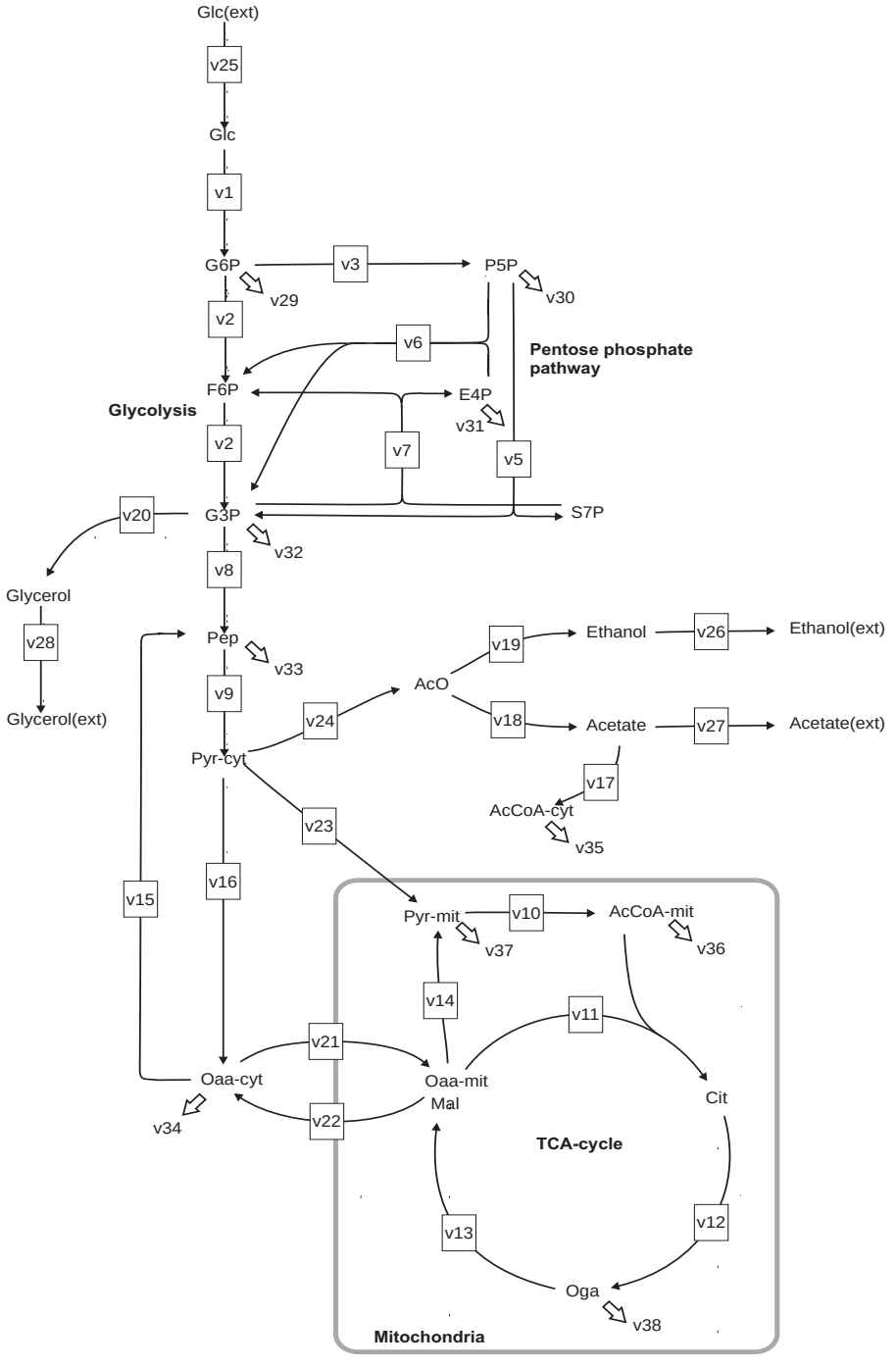

Figure 4 (Publication I). Metabolic network model of *S. cerevisiae* for net flux determination utilising flux ratios as additional constraints.

The flux ratio equations were set up according to the METAFoR analysis for the reactions in the stoichiometric models of the central carbon metabolism of *S. cerevisiae* (Equations 12 to 16, the reaction numbers are defined in Figure 4). Depending on the structure of the network of active reactions the flux ratio equations included one to six of the following (Equations 12 to 16):

the fraction of Pep from PPP assuming a maximal contribution of PPP

$$fr1 = \frac{v_5 + 3v_6 + 2v_7}{v_5 + 2v_4 + v_6} \tag{12}$$

the fraction of Pep originating from Oaa_{cyt}, X_{PEPck}:

$$fr2 = \frac{v_{15}}{v_{15} + v_8} \tag{13}$$

the fraction of Oaa_{mit} originating from Oaa_{cyt}, $X_{Oaa-transport}$:

$$fr3 = \frac{v_{21}}{v_{21} + v_{13}} \tag{14}$$

the fraction of Oaa_{cyt} originating from Pyr_{cyt}, $X_{Oaa_{cyt}\_from\_Pyr_{cyt}}$:

$$fr4 = \frac{v_{16}}{v_{16} + v_{22}} \tag{15}$$

the upper and lower bounds for Pyrmit originating from the malic enzyme reaction, X_{MAE_ub} and X_{MAE_lb}:

$$fr5 \le \frac{v_{14}}{v_{14} + v_9} \le fr6 \tag{16}$$

The following constraint equations were thus obtained from the flux ratio equations (Equations 17 to 22):

$$v_5 + 3v_6 + 2v_7 - fr1(v_5 + 2v_4 + v_6) = 0 \tag{17}$$

$$v_{15} - fr2(v_{15} + v_8) = 0 \tag{18}$$

$$v_{21} - fr3(v_{21} + v_{13}) = 0 \tag{19}$$

$$v_{16} - fr4(v_{16} + v_{22}) = 0 \quad (20)$$

$$v_{14} - fr5(v_{14} + v_9) = 0 \quad (21)$$

$$fr6(v_{14} + v_9) - v_{14} = 0 \quad (22)$$

Irreversibility was assumed for the intracellular fluxes v_3, v_4, v_8, v_9, v_{10}, x_{11}, v_{12}, v_{13}, v_{14}, v_{15}, v_{16}, v_{21}, v_{22}, v_{23}, v_{24}, for extracellular fluxes v_{25}, v_{26}, v_{27}, v_{28} and for the depletion of precursors to biosynthetic reactions and thus, only positive values were allowed for the fluxes. The minimization of the sum of the weighted square residuals of the metabolite mass balances was done using the Matlab function *fmincon* for optimization of constrained nonlinear multivariable function. The extracellular metabolite mass balances were assigned weights according to the experimental measurement error estimates. The biomass precursor metabolite mass balances were assigned ten-fold larger weights, relative to their stoichiometric coefficients in the biomass composition, since the assumption of constant biomass composition was expected to be harsh [Furukawa *et al.*, 1983]. The flux ratio constraints were included as strict constraints. The optimization was started 10000 times from random initial values to evaluate the uniqueness of the optimal flux distribution. The sensitivity of the flux distribution solutions to the noise in the flux ratio data and in the extracellular flux data was studied by Monte Carlo -simulations [Antoniewicz *et al.*, 2006]. Additive normally distributed noise with zero mean and experimentally determined variances of the flux ratios and the extracellular fluxes was sampled to the flux ratios and to the extracellular flux data, separately and simultaneously. A flux distribution was solved for each of the 100 sets of input data from 12 random initial flux distributions. Confidence intervals (95%) of the fluxes were determined.

4.8 Pathway reconstruction

In Publication III metabolic flux ratio (METAFoR) analysis was performed for *T. reesei* which lacks a genome-wide metabolic reconstruction. It was essential for the METAFoR analysis to reconstruct the biosynthetic pathways for proteinogenic amino acids of *T. reesei*. The reconstruction of amino acid biosynthetic pathways from their precursors in *T. reesei* was performed with ReTrace which is a recent computational carbon path analysis method [Pitkänen *et al.*, 2009], which can be queried to discover branching metabolic pathways in a uni-

versal metabolic database. ReTrace aims to find pathways which transfer as many atoms from source to target metabolites as possible.

The reaction database used in ReTrace analysis was KEGG LIGAND, retrieved in March 2009 [Kanehisa *et al.*, 2008]. Reaction database contained 7827 reactions, 15400 compounds. *Atom mappings*, that describe how atoms are transferred in a reaction from substrate to product metabolites, were defined for 33795 substrate-product pairs in the RPAIR database [Kotera *et al.*, 2004], which is a subdatabase of KEGG. All reactions were considered bidirectional. To compute reaction scores, a database consisting of 101136 sequences annotated with an EC number in UniProt version 9.3 [The UniProt Consortium, 2007] was queried with the 9129 protein sequences in *T. reesei* genome [Martinez *et al.*, 2008] by blastp [Altschul *et al.*, 1997] using e-value cutoff 10 to detect remote homologs. Each reaction in the KEGG database was assigned a score by taking the maximum BLAST score over all UniProt-*Trichoderma* sequence pairs, where UniProt sequence had been annotated with an EC number corresponding to the reaction. A total of 3974 reactions received a score in this procedure, while the remaining 3853 reactions were assigned a zero score. Reaction scores reflected the degree of evidence from the detection of sequence homology that there exists an enzyme catalyzing the reaction in *T. reesei*.

ReTrace operates on an atom-level graph representation of the metabolic network of all the reactions in the reaction database [Pitkänen *et al.*, 2009]. First, the metabolic network is converted into an atom graph, where nodes correspond to the atoms of metabolites and edges the atom mappings between the individual atoms. Atom mappings of carbon atoms were utilised in the reconstruction of the biosynthetic pathways of amino acids in *T. reesei*. Other atoms than carbons, such as nitrogen and sulphur, were not considered in the analysis performed in Publication III.

ReTrace [Pitkänen *et al.*, 2009] utilises a *K shortest paths* algorithm [Eppstein, 1994] to discover a number of connections between nodes in the atom graph. Given a query to find pathways from source to target metabolites (to amino acids in Publication III) ReTrace searched for shortest paths in the atom graph from any atom in source metabolites to any target metabolite atom [Pitkänen *et al.*, 2009]. Each shortest path was then processed. ReTrace traces back the target metabolite atoms along the atom mappings. Then branching points were identified as atom mappings that did not transfer the traced carbon atom towards the source [Pitkänen *et al.*, 2009]. Possible metabolites that could pro-

vide the missing carbon were determined and shortest paths from source to these metabolites were searched to determine the branches.

4.9 Localization of amino acid biosynthetic enzymes in *T. reesei*

In Publication III, TargetP, a machine learning method based on neural networks that predicts both chloroplast and mitochondrial targeting peptides and secretory signal peptides was utilized to predict the probable subcellular localization of some enzymes on the biosynthetic pathways of amino acids in *T. reesei* [Emanuelsson *et al.*, 2007; Emanuelsson *et al.*, 2000].

5. Results and discussion

In Publications I, II, III and IV ratios of intracellular fluxes were determined utilising ^{13}C-labelling experiments and the established metabolic flux ratio (METAFoR) analysis approach was extended. In Publication I flux ratios from METAFoR analysis combined with NMR spectroscopy as the analytical tool was utilized as experimentally derived additional constraints that enabled solving intracellular net fluxes under *S. cerevisiae* in different oxygenation conditions and determination of the dependence of the flux phenotype of *S. cerevisiae* on oxygen provision. In Publication II the Established METAFoR analysis was extended to a two-carbon source case to investigate the metabolism of *P. pastoris* under process conditions of recombinant protein production where methanol is used as inducer and glycerol as a carbon source. In Publication III the previously scarcely studied metabolism of filamentous fungus *T. reesei* was studied under different conditions of carbon catabolite repression. Since *T. reesei* lacks a genome-wide metabolic reconstruction the METAFoR analysis was coupled to essential reconstruction of the biosynthetic pathways of amino acids from genome level evidence. Comparison of the flux ratios of *T. reesei* to the ones observed in *S. cerevisiae* confirmed that the regulation of the central pathway fluxes is programmed in distinct ways in the two organisms. In Publication IV a systematic and analytic framework for derivation of flux ratio equations from a given model and ^{13}C-labelling data that constrains isotopomer distributions was developed. Comparison of the automatically derived flux ratios to manually by METAFoR analysis approach derived flux ratios was not straightforward because the biological information encoded in the METAFoR analysis approach by domain experts is diverse. When the model employed in the automatic derivation corresponded to the METAFoR approach, the derived flux ratios agreed well.

In Publication V NMR spectroscopy was utilized in verification of the reconstruction of a metabolic pathway in *T. reesei.* A reaction product of a step on the D-galacturonic acid catabolic pathway was identified by 1D and 2D NMR spectroscopy.

5.1 Utilization of ^{13}C-metabolic flux analysis excluded cofactor mass balances

In Publication I the dependence of intracellular flux distribution of *S. cerevisiae* on the level oxygenation was quantified by ^{13}C-MFA. *S. cerevisiae* was grown in glucose-limited chemostat cultures at a low dilution rate of 0.1 h^{-1} that ensured that the maximum respiratory rate was not reached even under fully aerobic conditions. The cultures were aerated with five different fractions of oxygen in the chemostat inlet gas: 20.9%, 2.8%, 1.0%, 0.5% and 0.0%. Wiebe *et al.* (2008) observed that at aeration of 20.9% O_2 in the chemostat inlet gas the metabolism of *S. cerevisiae* was fully respiratory and ethanol was observed in the medium under all the rest of the conditions. Duplicate cultures at each oxygenation condition were ^{13}C-labelled with uniform-labelling approach for ^{13}C-MFA analysis in Publication I. Quantitative ratios of merging fluxes in central carbon metabolism of *S. cerevisiae* were obtained by METAFoR analysis approach utilizing NMR spectroscopic detection of ^{13}C-labelling [Szyperski, 1995; Szyperski *et al.*, 1999; Maaheimo *et al.*, 2001]. The flux ratios were utilised as additional constraints to solve the mass balance equation system of the stoichiometric model of central carbon metabolism of *S. cerevisiae* as was earlier done for another eukaryote *P. anomala* with MS detected constraints by Fredlund *et al.* (2004) with the approach published by Fischer *et al.* (2004). ^{13}C-MFA enabled solving the intracellular net fluxes without including the redox cofactors in the mass balances which was essential since oxygen availability strongly affects the systems that maintain the redox homeostasis in cells.

5.2 Pyruvate branching point distribution responsive

The quantification of the net flux distributions of *S. cerevisiae* in response to different oxygenation conditions showed that the fluxes were redistributed not only between the cells grown in the fully aerobic conditions, under conditions of reduced oxygen provision and under anaerobic conditions but also between cells grown with different levels of low oxygen (2.8%, 1.0% and 0.5% O_2 in the

chemostat inlet gas) (Publication I). Although the metabolism of *S. cerevisiae* was respiro-fermentative under each of these low oxygen conditions, the actual amount of oxygen available resulted in different distribution to the respirative and fermentative pathways. The flux distribution at the pyruvate branch point, where the respirative and the fermentative pathways and the anaplerotic pathway, that operates to replenish the TCA cycle, diverge was particularly responsive to the level of reduction in oxygen provision. The respirative pathway flux decreased progressively under reduced oxygenation conditions where the availability of terminal electron acceptor limited the respiratory rate. However, the respiratory energy generation, that is highly efficient because of the high electronegativity of oxygen, provided a large fraction of ATP even under the low oxygen conditions (Table 1).

Table 1 (Publication I). Estimated fractions of respiration coupled ATP generation in *S. cerevisiae* under different oxygenation conditions.

	O_2 provided in the fermentor inlet gas				
	20.9%	2.8%	1.0%	0.5%	0.0%
ATP from respiration (%)	59	55	36	25	0
OUR (mmol $g^{-1}h^{-1}$)	2.7	2.5	1.7	1.2	0
ATP/ O[a]	0.9	1.0	1.1	1.1	-
ATP/ $2e^{-}$[b]	1.0	0.9	1.0	1.2	-

[a] Calculated from the oxygen uptake rate (OUR), [b]Calculated from the flux of electron donors to the respiratory chain.

5.3 Methanol and glycerol co-utilization extension

In Publication II a biosynthetically directed fractional ^{13}C-labelling approach was established for yeast *P. pastoris* growing on carbon substrate mixture of methanol and glycerol. Methanol is utilised as an inducer of protein production in *P. pastoris* processes. The approach allowed the quantification of the metabolic state of the TCA cycle and associated pathways under production conditions and thus was an important methodological expansion of the metabolic flux ratio (METAFoR) analysis [Szyperski, 1995; Szyperski *et al.*, 1999; Maaheimo *et al.*, 2001].

5.4 Flux distributions robust against different fractions of methanol

It was shown that co-assimilation of methanol as a carbon source does not alter the way the common amino acids are synthesized in *P. pastoris* growing on a sole multicarbon source, and that the growth on different glycerol/methanol mixtures at a given growth rate results in rather similar flux ratio profiles in the TCA cycle and related pathways as the fraction of methanol is increased (Publication II). In contrast, a clear effect of specific growth rate on the relative activity of the TCA cycle and related pathways was observed, regardless of the methanol fraction in the feed, consistent with the observation that TCA cycle activity in *S. cerevisiae* is strongly correlated with the environmentally determined specific growth rate [Blank and Sauer, 2004]. Co-assimilation of methanol as a carbon source has a clear impact with respect to the activity of the PPP, which is consistent with the increasing flux of methanol molecules towards the synthesis of central carbon metabolism intermediates (e.g. Pep), as observed when the fraction of methanol in the feed medium is increased. However, this pattern was not observed in cells growing at the higher dilution rate (where methanol is partially accumulated in the medium) suggesting that the distribution of methanol carbon into assimilatory and dissimilatory pathways may be different. Earlier ^{13}C-labelling studies of methanol metabolism of the methylotrophic yeast *H. polymorpha* [Jones and Bellion, 1991] showed that the linear methanol oxidation pathway to CO_2 only operates under extreme conditions (e.g. methanol accumulation to toxic levels), suggesting a role in detoxification.

The information from the ^{13}C-labelling and METAFoR analysis [Szyperski, 1995; Szyperski *et al.*, 1999; Maaheimo *et al.*, 2001] of *P. pastoris* on glycerol and methanol mixtures is valuable for the optimization of culture processes for the production of recombinant proteins in *P. pastoris*, where parameters such as the residual methanol concentration, specific growth rate, as well as mixed substrate culture strategies have been shown to have a dramatic impact on overall process productivity (Publication II). In addition, the information derived from the study may be relevant for the design of isotopic labelling experiments of recombinant proteins (or other cell components, e.g. cell wall glucans) for structural studies.

5.5 Metabolic flux ratio analysis of *T. reesei* necessitated reconstruction of biosynthetic pathways of amino acids

In Publication III the biosynthetic pathways of *T. reesei* were reconstructed for most of the proteinogenic amino acids with a computational carbon path analysis method ReTrace [Pitkänen *et al.*, 2009]. The method was used to search for pathways from a metabolic network consisting of all reactions found in a comprehensive metabolic reaction database, and to subsequently rank the pathways according to the degree of support from the *T. reesei*'s genome. Contiguous pathways, identical to the amino acid biosynthetic routes of *S. cerevisiae*, were found with high genome-level evidence. Origins of amino acids in *T. reesei* relevant for METAFoR analysis are shown in Figure 5.

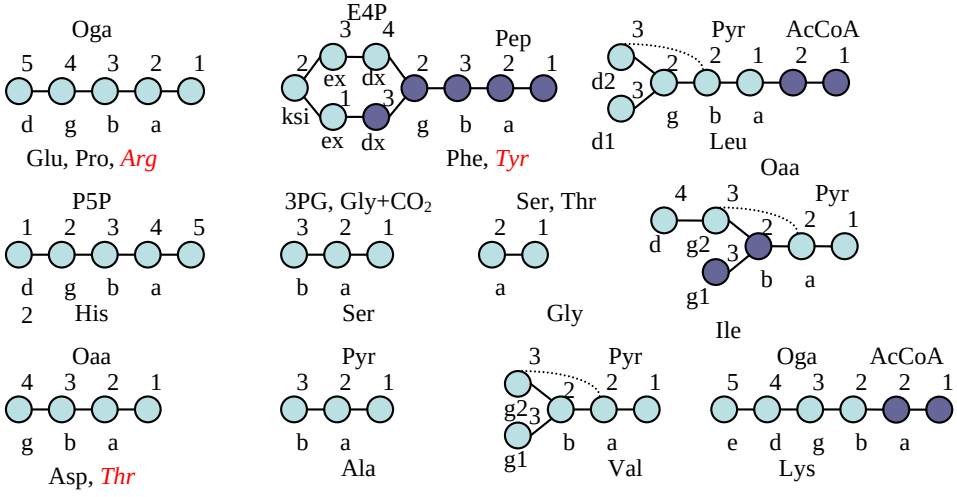

Figure 5 (Publication III). The origins of the carbon backbones of the proteinogenic amino acids in *T. reesei* that are relevant for METAFoR analysis and for which the biosynthetic pathways were reconstructed by computational pathway analysis method ReTrace [Pitkänen *et al.*, 2009]. If the biosynthetic pathway was not directly found by ReTrace, the amino acid abbreviation is denoted in red italics. The amino acid carbons are denoted in the following way: a = α, b = β, g = γ, d = δ, e = ε, ksi = ξ.

5.6 Primary respiratory metabolism

T. reesei wild type and Δ*cre1* strains were ^{13}C-labelled with uniform ^{13}C-labelling approach in batch cultures in flasks on repressive carbon source glucose and on sorbitol. The ^{13}C-labelling patterns of proteinogenic amino acids were in good

accordance with the compartmentalized model of eukaryotic central carbon metabolism, originally developed for *S. cerevisiae* [Maaheimo *et al.*, 2001]]. However, in contrast to *S. cerevisiae*, in both *T. reesei* strains Asp synthesis was observed to occur primarily from the mitochondrial pool of Oaa, under all the studied conditions.

The *T. reesei* wild type strain is known to exhibit carbon catabolite repression of hydrolytic gene expression during growth on glucose, whereas in the Δ*cre1* strain the repression is partially disturbed [Ilmén *et al.*, 1996]. The respiratory pathway does not become transcriptionally downregulated by the carbon catabolite repression in *T. reesei* as in *S. cerevisiae* [Chambergo *et al.*, 2002]. However, it is the *in vivo* fluxes that are the ultimate phenotype of an organism. The ratios of *in vivo* fluxes of *T. reesei* were solved with METAFoR analysis from the ^{13}C-labelling patterns of proteinogenic amino acids [Szyperski, 1995; Szyperski *et al.*, 1999; Maaheimo *et al.*, 2001] in the different conditions of carbon catabolite repression. This was the first time that the effect of carbon catabolite repression *T. reesei* on *in vivo* fluxes was quantitatively studied.

The relative anaplerotic flux, the flux that replenishes the TCA cycle, compared to the respiratory pathway flux was characteristic to primarily respiratory metabolism in the both *T. reesei* strains under all the studied conditions (Table 2). This indicated that *T. reesei* utilizes primarily respiratory metabolism also on preferred carbon source glucose. However, the observed relative anaplerotic fluxes suggested that the respirative activity of the TCA cycle in *T. reesei* was even slightly higher on the neutral carbon source sorbitol than on glucose. Only minor differences were observed between the *in vivo* flux distributions of the wild type and the Δ*cre1* strains. Therefore, CRE1 the key repressor of utilization of alternative carbon sources, does not mediate carbon source dependent metabolic state alterations in central carbon metabolism in *T. reesei*. The sophorose induction of cellulase gene expression did not result in significant changes in the relative requirements of proteinogenic amino acids or in the ratio of anabolic to oxidative activities of the TCA cycle.

Table 2 (Publication III). Comparison of Metabolic Flux Ratio (METAFoR) analysis of *T. reesei* to the crabtree positive and negative yeasts *S. cerevisiae* and *P. stipitis*. METAFoR analysis from fractional ^{13}C-labelling of *T. reesei* wild type and *Δcre1* strains in aerobic batch cultures on glucose and on sorbitol (Publication III), compared with the ones observed in *S. cerevisiae* and *P. stipitis* in aerobic batch cultures on glucose and in aerobic glucose-limited chemostat cultures [Maaheimo *et al.*, 2001; Fiaux *et al.*, 2003].

organism	*T. reesei*		*T. reesei*		*T. reesei*		*T. reesei*		*S. cerevisiae*		*S. cerevisiae*		*P.stipitis*		*P.stipitis*	
strain	wild type		delta *cre1*		wild type		delta *cre1*									
carbon source	glucose		glucose		sorbitol		sorbitol		glucose		glucose		glucose		glucose	
culture	batch		batch		batch		batch		batch		chemostat		batch		chemostat	
reference									Maaheimo		Fiaux		Fiaux		Fiaux	
	%	sd	%	sd	%	sd	%	sd	%	sd	%	sd	%	sd	%	sd
Pep from PPP (UB, no PEPck)	39	2	47	4	36	7	45	9	0-4		40	8	57	9	61	11
R5P from T3P and S7P	51	1	42	1	72	3	79	4	68	2	59	2	57	2	72	2
R5P from E4P	25	2	23	1	46	2	54	3	10	2	33	2	35	2	43	2
Anaplerotic flux ratio	35	1	33	2	26	3	42	5	76	4	31	2	36	2	32	2
MAE (UB)	4	0	9	1	12	2	6	5	25-30		<13		<6		<7	
MAE (LB)	2	0	6	1	9	1	4	3	nd	nd	nd	nd	nd	nd	nd	nd

UB upper bound, LB lower bound, nd not determined

5.7 Framework for analytic and systematic derivation of flux ratio equations

In Publication VI a systematic and analytic framework for ^{13}C-metabolic flux ratio analysis was introduced. Previously the utilisation of the METAFoR analysis [Szyperski, 1995; Szyperski *et al.*, 1999; Maaheimo *et al.*, 2001] has been relying on manual derivation of the equations that constrain the flux ratios. The systematic and analytic framework for ^{13}C-metabolic flux ratio analysis is a generalization and formalization of existing analytic methods for computing metabolic flux ratios [Maaheimo *et al.*, 2001; Zamboni *et al.*, 2005; Szyperski 1998] and facilitates an efficient and analytic computation of the ratios between the fluxes producing the same junction metabolite in a given metabolic network.

The model of the central carbon metabolism of *S. cerevisiae* was formulated also in a ^{13}C-FLUX [Wiechert *et al.*, 2001] format and artificial substrate labelling was employed to obtain simulated ^{13}C-labelling data for verification of the implementation of the framework for flux ratio determination. Then NMR spectroscopy data from ^{1}H-^{13}C HSQC experiments, relative intensities of fine structures that represent different combinations of ^{13}C and ^{12}C atoms coupled to a central ^{13}C atom in proteinogenic amino acids, was utilized to compare the flux ratios derived with the implemented framework with manually derived flux ratios. The experiment showed that the framework was able to provide relevant quantitative information on the distribution of metabolic fluxes, even when only constraints to the isotopomer distributions of proteinogenic amino acids are

measured. Detection of the ^{13}C-labelling status of 15 proteinogenic amino acids resulted in flux ratios for junctions in four cytosolic metabolites (Oaa, Pep, Gly and Ser) and in three mitochondrial metabolites (Oaa, AcCoA and Pyr). In addition, an upper bound for G3P molecules that had gone through a transketolase reaction could be solved when the model was first manually simplified to correspond to the model utilised in the manual derivation of the G3P flux ratio. The computed flux ratios where compared to the manually derived ones whenever the assumptions made in the manual derivation were consistent with the general model employed in the automatic derivation of the flux ratios. The automatically derived flux ratios agreed well with the manually derived ratios. Differences between the estimations could be explained by numerical instabilities and by differences in computational procedures: in manually derived ratios the estimations are based on the breakage of a single bond in different routes leading to a metabolite while in the developed framework more isotopomer information is possibly utilized in the estimation.

5.8 L-threo-3-deoxyhexulosonate is a reaction product of L-galactonate dehydratase

In Publication V the metabolism of fungus *T. reesei* was further studied and an L-galactonate dehydratase and the corresponding gene were identified from *T. reesei*. The enzyme converts L-galactonate to L-threo-3-deoxy-hexulosonate (2-keto-3-deoxy- L-galactonate) and belongs to the fungal pathway for D-galacturonic acid catabolism. L-galactonate dehydratase is the second enzyme of the pathway after the D-galacturonic acid reductase. L-galactonate dehydratase activity is present in *T. reesei* mycelia grown on D-galacturonic acid but absent when other carbon sources are used for growth. L-galactonate dehydratase is active on sugars L-galactonate and D-arabonate in which the hydroxyl groups of the C2 and the C3 in the Fischer projection are in L- and D-configuration, respectively. The enzyme was not active with sugar acids having the hydroxyl groups of C2 in D-configuration and C3 in L-configuration as in D-galactonate, D-gluconate and D-xylonate and with sugar acids having the hydroxyl groups of C2 and C3 in D-configuration as in D-gulonate.

In order to define the reaction product of L-galactonate dehydratase it was analysed by NMR spectroscopy. To generate a sufficient amount of reaction product L-galactonate was incubated in the yeast extract of the strain expressing the L-galactonate dehydratase gene. In this extract the reaction product did not

react further, which facilitated the NMR spectroscopic analysis. In the *T. reesei* mycelia extract the reaction product was degraded, making the NMR spectroscopic analysis more difficult. The NMR spectroscopic analysis showed that erythro- or threo-3-deoxy-hexulosonate was formed. Knowing the substrate of the dehydratase reaction it was concluded that the product was L-threo-3-deoxyhexulosonate. The NMR spectroscopic analysis also revealed that it was predominantly in the pyranose form. For the pyranose form two anomers are possible; the carboxyl group in R_1 and the hydroxyl group in R_2 or vice versa. The NMR spectroscopic analysis suggested that one of the anomers was predominant but it did not allow determination of which of the two anomers it was. The NMR spectroscopic analysis revealed also that the axial hydrogen at C3 is the hydrogen that was added in the reaction. However, as there are two possible chair conformations of the pyranose ring it remained unclear which of the two protons is in the axial position.

6. Conclusions and prospects

The analyses of flux ratio of different organisms revealed phenotypic differences in detail. In the studies of this thesis the differences between the regulatory principles of two eukaryotic organisms, *S. cerevisiae* and *T. reesei*, were observed on flux phenotypic level (Publications I and III). The former pursues high rate of ATP production whereas the latter seeks a high ATP yield. The distributions of fluxes to respirative and fermentative pathways were similar when *S. cerevisiae* was growing slowly with unlimited oxygenation and low glucose and *T. reesei* was growing on high glucose at maximum rate. However, the fast growing *S. cerevisiae* on high glucose diminishes the respirative pathway flux and speeds up the glycolytic flux and fermentative pathway activity [De Deken, 1966; Maaheimo *et al.* 2001; Gombert *et al.* 2001]. The differences in regulatory principles of the two organisms can be explained by the different natural habitats of the organisms and adaptive evolution of the regulatory systems. Pfeiffer *et al.* (2001) have further claimed that the preference to high ATP yield has contributed to the development of multicellular organisms.

6.1 Robust regulatory system enables stable flux phenotype

Complex and multi-level regulatory mechanisms can maintain fairly stable distribution of fluxes in altered conditions. However, attenuating changes can though be observed in the underlying levels of transcription, proteome and metabolome [Davies and Brindle, 1992; Schaaff *et al.*, 2004; Raamsdonk *et al.*, 2001]. Furthermore, an adaptation of the flux phenotype to altered conditions can occur through sequential changes in the underlying levels of cell function [van der Brink *et al.*, 2008]. There is interdependence between metabolite concentrations, enzyme concentrations and fluxes through the metabolic network,

enzyme kinetics and regulation of enzyme activity [Stryer, 1995]. Enzyme concentration is set by hierarchical regulatory system that is closely linked to the metabolic status of an organism. Nutrient sensing and feedback regulation trigger signalling cascades that affect the hierarchical regulatory system [Zaman *et al.*, 2008; Zaman *et al.*, 2009]. Coordinated responses have been observed even between metabolites and transcripts [Bradley *et al.*, 2009]. The enzyme activity cannot affect the equilibrium constant of a chemical reaction but the network can enable a shift of an effective equilibrium constant through futile cycles [Qian and Beard, 2006]. The multilevel regulatory system provides robustness that enables stability of flux phenotypes. Are the robust response mechanisms to different perturbations mechanistically similar? How has the robustness of an organism against fluctuations in its natural environment developed? The same multilevel regulatory system provides fine-tuned adaptation in some conditions. A fine piece of work was published by Bennet and co-workers (2008) where they could conclude that the regulatory system of galactose metabolism in *S. cerevisiae* functions as a low-pass filter that in dynamic conditions enables adaptation to slow changes and robustness against fast perturbations. Thus the same mechanisms can provide both robustness and adaptation. Are they different depending on the perturbation? How are the decisions between robust response and adaptation made when the cells are exposed to different perturbations? These are interesting questions that are expected to be answered by sophisticated systems biology studies when the experimental and modelling tools are reaching an adequate performance level.

It should eventually be possible to integrate data of all measurable cell constituents and integrate it with models of cellular interaction networks and compose a predictive model of cell function. There are methods for recording data on all the effector types. Quantitative metabolomics methods in particular by MS have been established [Ewald *et al.*, 2009; Buscher *et al.*, 2009]. Enzyme concentrations are determined by rates of transcription, mRNA degradation, translation and protein degradation and both protein and transcript levels can be observed in large-scale [de Groot *et al.*, 2007]. Regulatory mechanisms that affect fluxes are for example enzyme phosphorylations which can currently be monitored [Ptacek *et al.*, 2005; Huber *et al.*, 2009]. Information on allosteric interactions has been reported in literature and some is collected for example in database BRENDA (http://www.brenda-enzymes.org) [Schomburg *et al.*, 2002]. All effectors are tied together by interaction networks that are dynamic in reality. However, the static interaction network models provide a framework for the

dynamic system responses that give rise to the observed flux phenotypes. Hindrances can still be pinpointed both in experimental methods and in modelling tools. Not all the levels of cell function can be monitored in genome-wide scale and modelling tools need to be able to handle that missing data and uncertainties as well as to find suitable means to model different types of interactions and information transfer through them. Ishii *et al.* (2007) performed a pioneering data integration study in a limited system were flux phenotypic data and data on underlying regulatory levels was simultaneously recorded from central carbon metabolism of single gene deletion mutant strains of *E. coli* under different growth conditions. However, since the scale of the system was limited to the central carbon metabolism, integrative visualisation of the metabolic network was adequate for interpretation of the data. Figure 6 shows an example of integrative data analysis of the switch from fully respiratory to respiro-fermentative phenotype of *S. cerevisiae* in context of interaction network of limited extent (unpublished results). The systemic co-responses are almost too complex to visually comprehend even though the system is limited. Multiomics data integration requires tools that generate hypotheses of systemic response mechanisms.

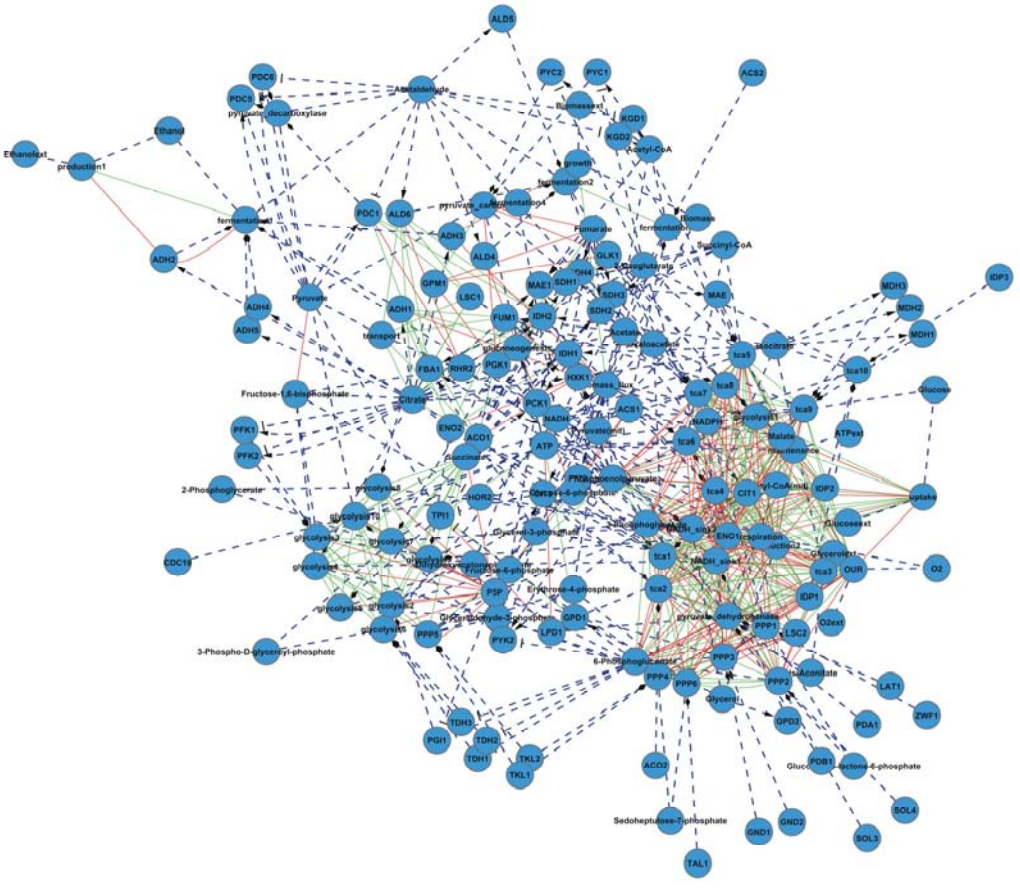

Figure 6. Co-response interactions in *S. cerevisiae* between components of central carbon metabolism in switch from fully respirative to respiro-fermentative metabolism in limited oxygen provision (unpublished results). Co-responses are calculated according to Raamsdonk *et al.* (2001) as ratios of the log fold changes in the quantities of components between the two conditions [Publication I; Rintala *et al.*, 2009]. Co-response interactions are drawn for the components having strong positive (in green) or negative (in red) co-responses.

6.2 Determinants of energy generation processes

The biological features studied in the Publications included in the thesis were effects of oxygen and different carbon sources on flux distributions of three eukaryotes *S. cerevisiae*, *T. reesei* and *P. pastoris*. Carbon sources among other nutrients [Rohde *et al.*, 2008] are known to trigger regulatory cascades that can ultimately determine the active energy production processes as in case of high

glucose sensed by *S. cerevisiae* [Zaman *et al.*, 2009]. Oxygen on the other hand is a major determinant of active energy processes in aerobic organisms. Regulation of energy processes in bioproduction hosts is naturally of importance for developed bioprocess control. Many of the regulatory mechanisms are conserved among eukaryotes and thus, the studies of simple eukaryotes such as yeasts and fungi provide a basis also for understanding the function of higher cells. Understanding of the regulation of energy generation processes is central in investigations of for example human metabolic diseases. One of the metabolic diseases, type II diabetes, is at present increasing in population of western countries [Wild *et al.*, 2004]. Energy generation processes are central also in sports. Oxygen uptake rate and carbohydrate refuelling are central parameters of sports performance [Hulston and Jeukendrup, 2009]. Oxygen provides not only an electron acceptor for efficient energy generation but also a threat of cell component damage as a strong oxidant. Both features of oxygen are related to cell ageing and thus, of major interest [Koc *et al.*, 2004; Oliveira *et al.*, 2008; Finkel and Holbrook, 2000; Lin *et al.*, 2002].

6.3 Prospects of local flux ratio analysis

Local flux ratio analysis is an efficient approach for quantitative profiling of the flux phenotype. The conventional method is extendable from inferring ^{13}C-labelling patterns of proteinogenic amino acids to detection of free metabolites and from established network structures to novel or engineered organisms by systematic derivation of constraint equations (Publication IV). The strength and advantage of the local flux ratio analysis is the independence of external fluxes and therefore, also an independence of the definition of biomass effluxes [Zamboni *et al.*, 2009]. Biomass effluxes are usually derived from growth rate and biomass composition and the experimental determination of the detailed macromolecular composition of biomass is laborious [Lange 2002; Lange and Heijnen, 2001]. Thus, flux ratios provide independent quantitative measures of distribution of fluxes also when the effluxes from the system to macromolecule synthesis is not exactly known. Analysis can be targeted to metabolic junctions of interest or as many junctions can be analysed as possible to gain constraints for net flux determination by ^{13}C-MFA.

6.4 Large-scale flux analysis

Global iterative fitting becomes computationally unfeasible when the number of isotopomer balance equations explodes. System extension requires also additional measurements of ^{13}C-labelling patterns to gain constraints for isotopomer distributions of added metabolites. Without enough constraints the flux distribution cannot be accurately solved. Combinations of ^{13}C-labelling based constraints and an objective function have been proposed for example by Blank *et al.* (2005). Nevertheless, the large-scale models can always be utilised as scaffolds where the active networks can be identified and more detailed models built [Suthers *et al.*, 2007].

Flux balance analysis (FBA) is feasible in well-defined genome scale metabolic networks. The difficulty of definition of a relevant objective function limits the utilisation of FBA in flux determination. In determination of metabolic capabilities of organisms it is highly efficient. Automatic means to infer or to identify objective functions have been proposed [Knorr *et al.* 2007; Gianchandani *et al.* 2008]. As interest in systems biology is also on higher cells, intelligent definitions of objective functions have been proposed [Heuett and Qian, 2006] and are under development.

6.5 Flux analysis in dynamic conditions

Flux analysis is turning dynamic since stationary metabolic flux analysis is not convenient for those time-dependent processes that are of biotechnological interest or compatible with analysis of fluxes in higher cells. If the specific growth rate is low, prolonged labelling time is required to reach an isotopic pseudo steady-state in macromolecule components. During prolonged labelling time the steady state may alter. Furthermore, the non-growing cells are completely incompatible with stationary biosynthetically directed ^{13}C-labelling. Direct measurement of the labelling patterns of metabolic intermediates was thought to shorten the required labelling times but the investigations showed that the labelling patterns of intermediates stabilize much later than expected. Large pools of storage compounds, protein turnover and compartmentalized pools increase the stabilisation times of the labelling patterns of metabolic intermediates to the timescale of label stabilisation in macromolecule components [Aboka *et al.*, 2009; Grotkjaer *et al.*, 2004; den Hollander *et al.*, 1981].

Dynamic ^{13}C-MFA that significantly shortens the required labelling times emerged first to metabolic stationary but isotopically non-stationary conditions [Nöh *et al.*, 2007; Selivanov *et al.*, 2006; Schaub *et al.*, 2008; Hoffmann *et al.*, 2008]. Due to the shorter labelling times, the dynamic ^{13}C-flux analysis has broader applicability to mammalian cell cultures than the conventional stationary flux analysis. Mammalian cells grow usually slower than microbial cells and steady states are hard to sustain. Recently Munger *et al.* (2008) utilised dynamic ^{13}C-labelling of cultured human fibroblasts to determine metabolic targets for antiviral therapy. Extension of ^{13}C-MFA to metabolic non-stationary states was shown *in silico* by Wahl *et al.* (2008) and applied already to *E. coli* cultures by Schaub *et al.* (2008), and to hepatic cells in a two-part study [Hoffmann *et al.*, 2008; Maier *et al.*, 2008]. ^{13}C-MFA in transient state takes advantage of a kinetic model of the metabolic system, measurements of metabolic pool sizes and time-dependent labelling patterns. Integration of data yields different things depending on the approach. The kinetic parameters can be determined more reliably, non-measured metabolite pools can be solved and sensitivities of flux distribution solutions can be decreased by integrating stationary and non-stationary data [Nöh *et al.*, 2007; Selivanov *et al.*, 2006; Schaub *et al.*, 2008]. Dynamic ^{13}C-flux analysis is also suitable for stimulus-response experiments were the systemic time-dependent responses to perturbations are investigated [Wahl *et al.*, 2008]. ^{13}C-labelling experiments in dynamic conditions enable direct probing of robustness and control of the metabolic system and the data also reveals compartmentalization of metabolite pools [Schryer *et al.*, 2009]. Furthermore, the short time-scale ^{13}C-labelling reduces the cost of experiments and thus, is compatible with high-throughput experiments.

References

Aboka FO, Heijnen JJ, van Winden WA: Dynamic ^{13}C-tracer study of storage carbohydrate pools in aerobic glucose-limited *Saccharomyces cerevisiae* confirms a rapid steady-state turnover and fast mobilization during a modest stepup in the glucose uptake rate. *FEMS Yeast Res.* 2009, 9:191–201.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997, 25:3389–3402.

Antoniewicz MR, Kelleher JK, Stephanopoulos G: Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab Eng.* 2006, 8:324–337.

Antoniewicz MR, Kelleher JK, Stephanopoulos G: Elementary metabolite units (EMU): A novel framework for modeling isotopic distributions. *Metab Eng.* 2007, 9:68–86.

Araúzo-Bravo M, Shimizu K: An improved method for statistical analysis of metabolic flux analysis using isotomomer mapping matrices with analytical expressions. *J Biotechnol.* 2003, 105:117–133.

Arava Y, Wang Y, Storey JD, Liu CL, Brown PO, Herschlaq D: Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *PNAS*, 2003, 100:3889–3894.

Arita M: In Silico Atomic Tracing by Substrate-Product Relationships in *Escherichia coli* Intermediary Metabolism. *Genome Res.* 2003, 13:2455–2466.

Aro N, Pakula T, Penttilä M: Transcriptional regulation of plant cell wall degradation by filamentous fungi. *FEMS Microbiol Rev.* 2005, 29:719–735.

Bakker BM, Overkamp KM, van Maris AJA, Kötter P, Luttik MAH, van Dijken JP, Pronk JT: Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. *FEMS Microbiol Rev.* 2001, 25:15–37.

Beard DA, Babson E, Curtis E, Qian H: Thermodynamic constraints for biochemical networks. *J Theor Biol*. 2004, 228:327–333.

Beard DA, Liang SD, Qian H: Energy balance for analysis of complex metabolic networks. *Biophys J.* 2002, **83**:79–86.

Beard DA, Qian H: Thermodynamics-based computational profiling of cellular regulatory control in hepatocyte metabolism. *Am J Physiol Endocrino Metab.* 2005, 288:E633–E644.

Becker SA, Palsson BO: Context-specific metabolic networks are consistent with experiments. *PLOS Comput Biol.* 2008, 4: e1000082.

Bennet MR, Lee Pang W, Ostroff NA, Baumgartner BL, Nayak S, Tsimring LS, Hasty J: Metabolic gene regulation in a dynamically changing environment. *Nature* 2008, 454:1119–1122.

Bernhard SA, Tompa P: The mechanism of succinate or fumarate transfer in the tricarboxylic acid cycle allows molecular rotation of the intermediate. *Arch Biochem Biophys*. 1990, 276:191–198.

Blank LM, Kuepfer L, Sauer U: Large-scale ^{13}C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol.* 2005, 6:R49.

Blank LM, Sauer U: TCA cycle activity in *Saccharomyces cerevisiae* is a function of the environmentally determined specific growth and glucose uptake rates. *Microbiology* 2004, 150:1085–1093.

Boles E, de Jong-Gubbels P, Pronk JT: Identification and Characterization of *MAE1*, the *Saccharomyces cerevisiae* Structural Gene Encoding Mitochondrial Malic Enzyme. *J Bacteriol.* 1998, 180:2875–2882.

Botstein D, Chervitz SA, Cherry JM: Yeast as a model organism. *Science* 1997, 277:1259–1260.

Bradley PH, Brauer MJ, Rabinowitz JD, Troyanskaya OG: Coordinated concentration changes of transcripts and metabolites in *Saccharomyces cerevisiae*. *PLOS Comput Biol*. 2009, 5:e1000270.

Brockmann R, Beyer A, Heinisch JJ, Wilhelm T: Posttranscriptional expression regulation: what determines translation rates? *PLoS Comput Biol.* 2007, 3:e57.

Brosnan JT, Brosnan ME, Nissim I: A window into cellular metabolism: hepatic metabolism of ^{15}N-labelled substrates. *Metab Eng.* 2004, 6:6–11.

Burgard AP, Maranas CD: Optimization-Based Framework for Inferring and Testing Hypothesized Metabolic Objective Functions. *Biotechnol Bioeng*. 2003, 82:670–677.

Burgard AP, Pharkya P, Maranas CD: OptKnock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng*. 2003, 85:1–11.

Buscher JM, Czernik D, Ewald JC, Sauer U, Zamboni N: Cross-platform comparison of methods for quantitative metabolomics of primary metabolism. *Anal Chem.* 2009, 81:2135–2143.

Butow RA, Avadhani NG: Mitochondrial signalling: the retrograde response. *Mol Cell* 2004, 14:1–15.

Castello PR, David PS, McClure T, Crook Z, Poyton RO: Mitochondrial cytochrome oxidase produces nitric oxide under hypoxic conditions: Implications for oxygen sensing and hypoxic signalling in eukaryotes. *Cell Metab*. 2006, 3:277–287.

Chambergo FS, Bonaccorsi ED, Ferreira AJS, Ramos ASP, Ribamar Ferreira Júnior J, Abrahao-Neto J, Farah JPS, El-Dorry H: Elucidation of the Metabolic Fate of Glucose in the Filamentous Fungus *Trichoderma reesei* Using Expressed Sequence Tag (EST) Analysis and cDNA Microarrays. *J Biol Chem.* 2002, 277:13983–13988.

Chen RE, Thorner J: Function and regulation in MAPK signaling pathways: Lessons learned from the yeast *Saccharomyces cerevisiae*. *Biochim Biophys Acta* 2007, 1773:1311–1340.

Christensen B, Nielsen J: Isotopomer analysis using GC-MS. *Metab Eng.* 1999, 1:282–290.

Cortassa S, Aon MA: The onset of fermentative metabolism in continuous cultures depends on the catabolite repression properties of *Saccharomyces cerevisiae. Enzyme Microb Technol*. 1998, 22:705–712.

Covert MW and Palsson BO: Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J Biol Chem* 2002, 277:28058–28064.

Croasmun WR, Carlson RMK: *Two-dimensional NMR spectroscopy: Applications for Chemists and Biochemists*. 2nd edition, VHC Publishers, Inc., New York, 1994, 958 p.

Cziferszky A, Mach RL, Kubicek CP: Phosphorylation Positively Regulates DNA Binding of the Carbon Catabolite Repressor Cre1 of *Hypocrea jecorina (Trichoderma reesei)*. *J Biol Chem.* 2002, 277:14688–14694.

Daran-Lapujade P, Rossell S, van Gulik WM, Luttik MAH, de Groot MJL, Slijper M, Heck AJR, Daran J-M, de Winde JH, Westerhoff HV, Pronk JT, Bakker BM: The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. *PNAS* 2007, 104:15753–15758.

D'Autréaux B, Toledano MB: ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat rev Mol Cell Biol*. 2007, 8:813–824.

Davies BS, Rine J: A role for sterol levels in oxygen sensing in *Saccharomyces cerevisiae*. *Genetics* 2006, 174:191–201.

Davies SE, Brindle KM: Effects of overexpression of phosphofructokinase on glycolysis in the yeast *Saccharomyces cerevisiae*. *Biochemistry* 1992, 19:4729-35.

De Deken RH: The Crabtree effect: a regulatory system in yeast. *J Gen Microbiol.* 1966, 44:149–156.

de Groot MJL, Daran-Lapujade P, van Breukelen B, Knijnenburg TA, de Hulster EAF, Reinders MJT, Pronk JT, Heck AJR, Slijper M: Quantitative proteomics and transcriptomics of anaerobic and aerobic yeast cultures reveals post-transcriptional regulation of key cellular processes. *Microbiology* 2007, 153:3864–3878.

de Jong-Gubbels P, Bauer J, Niederberger P, Stückrath I, Kötter P, van Dijken JP, Pronk JT: Physiological characterisation of a pyruvate-carboxylase-negative *Saccharomyces cerevisiae* mutant in batch and chemostat. *Antonie Van Leeuwenhoek* 1998, 74:253–263.

Den Hollander JA, Behar KI, Shulman RG: ^{13}C NMR study of transamination during acetate utilization by *Saccharomyces cerevisiae*. *PNAS* 1981, 78:2693–2697.

Dirmeier R, O'Brien KM, Engle M, Dodd A, Spears E, Poyton RO: Exposure of yeast cells to anoxia induces transient oxidative stress: Implications for the induction of hypoxic genes. *J Biol Chem.* 2002, 277:34773–34784.

Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, Srivas R, Palsson BO: Global reconstruction of the human metabolic network based on genomic and bibliomic data. *PNAS* 2007, 104, 1777–1782.

Emanuelsson O, Brunak S, von Heijnen G, Nielsen H: Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc.* 2007, 2:953–971.

Emanuelsson O, Nielsen H, Brunak B, von Heijnen G: Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol.* 2000, 300:1005–1016.

Eppstein D: Finding the k shortest paths. *35th IEEE Symp. Foundations of Comp. Sci*., Santa Fe, 1994, pp. 154–165.

Ewald JC, Heux S, Zamboni N: High-Throughput Quantitative Metabolomics: Workflow for Cultivation, Quenching, and Analysis of Yeast in a Multiwell Format. *Anal Chem.* 2009, 81:3623–3629.

Feist AM, Herrgård MJ, Thiele I, Reed JL, Palsson BO: Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol*. 2009, 7:129–143.

Fiaux J, Cakar ZP, Sonderegger M, Wutrich K, Szyperski T, Sauer U: Metabolic-flux profiling of the yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*. *Eukaryot Cell* 2003, 2:170–180.

Finkel T, Holbrook NJ: Oxidants, oxidative stress and the biology of ageing. *Nature* 2000, 408:239–247.

Fischer E, Sauer U: Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS. *Eur J Biochem.* 2003, 270:880–891.

Fischer E, Zamboni N, Sauer U: High-throughput metabolic flux analysis based on gas chromatography–mass spectrometry derived ^{13}C constraints. *Anal Biochem*. 2004, 325:308–316.

Francke C, Siezen RJ, Teusink B: Reconstructing the metabolic network of a bacterium from its genome. *Trends Microbiol.* 2005, 13:550–558.

Fredlund E, Blank LM, Schnurer J, Sauer U, Passoth V: Oxygen- and glucose-dependent regulation of central carbon metabolism in *Pichia anomala*. *Appl Env Microbiol.* 2004, 70:5905–5911.

Frick O, Wittmann C: Characterization of the metabolic shift between oxidative and fermentative growth in *Saccharomyces cerevisiae* by comparative ^{13}C flux analysis. *Microb Cell Fact*. 2005, 4:30.

Friebolin H: *Basic One- and Two-dimensional NMR Spectroscopy*. VCH Verlagsgesellschaft mbH, D-6940 Weinheim 1991, 344p.

Fuda NJ, Ardehali MB, Lis JT: Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature*, 2009, 461:186–192.

Furukawa K, Heinzle E, Dunn IJ: Influence of Oxygen on the Growth of *Saccharomyces cerevisiae* in Continuous Culture. *Biotechnol Bioeng.* 1983, 25:2293–2317.

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

Gadian DG: *Nuclear magnetic resonance and its applications to living systems*. Oxford University Press, Oxford 1982, 195p.

Gancedo JM: Yeast carbon catabolite repression. *Microbiol Mol Biol Rev.* 1998, 62:334–361.

Ghosh S, Zhu T, Grossmann I, Ataai M, Domach M: Closing the loop between feasible flux scenario identification for construct evaluation and resolution of realized fluxes via NMR. *Computers and Chemical Engineering* 2005, 29:459–466.

Gianchandani EP, Oberhardt MA, Burgard AP, Maranas CD, Papin JA: Predicting biological system objectives de novo from internal state measurements. *BMC Bioinformatics* 2008, 9:43.

Gombert AK, dos Santos MM, Christensen B, Nielsen J: Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol.* 2001, 183:1441–1451.

Grotkjaer T, Akesson M, Christensen B, Gombert AK, Nielsen J: Impact of transamination reactions and protein turnover on labeling dynamics in (13)C-labeling experiments. *Biotechnol Bioeng.* 2004, 86:209–216.

Heijnen JJ: Approximative kinetic formats used in metabolic network modelling. *Biotechnol Bioeng*. 2005, 91:534–545.

Herrero E, Ros J, Bellí G, Cabiscol E: Redox control and oxidative stress in yeast cells. *Biochim Biophys Acta* 2008, 1780:1217–1235.

Herrgård MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Blüthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novère N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasić I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kirdar B, Penttilä M, Klipp E, Palsson BØ, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB: A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat Biotechnol.* 2008, 26:1155–1160.

Heuett WJ, Qian H: Combining flux and energy balance analysis to model large-scale biochemical networks. *J Bioinform Comput Biol*. 2006, 4:1227–1243.

Hoffmann U, Maier K, Niebel A, Vacun G, Reuss M, Mauch K: Identification of metabolic fluxes in hepatic cells from transient ^{13}C-labelling experiments: Part I. Experimental observations. *Biotechnol Bioeng.* 2008, 100:344–354.

Holzhutter H-G: The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. *Eur J Biochem.* 2004, 271:2905–2922.

Hon T, Dodd A, Dirmeier R, Gorman N, Sinclair PR, Zhang L, Poyton RO: A mechanism of oxygen sensing in yeast: multiple oxygen-responsive steps in the heme biosynthetic pathway affect Hap1 activity. *JBC* 2003, 278:50771–50780.

Huber A, Bodenmiller B, Uotila A, Stahl M, Wanka S, Gerrits B, Aebersold R, Loewith R: Characterization of the rapamycinsensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis. *Genes Dev.* 2009, 23:1929–1943.

Hulston CJ, Jeukendrup AE: No placebo effect from carbohydrate intake during prolonged exercise. *Int J Sport Nutr Exerc Metab.* 2009, 19:275–284.

Ilmén M: *Molecular mechanisms of glucose repression in the filamentous fungus Trichoderma reesei*. PhD thesis. VTT Publications 315, VTT Espoo; 1997. 86 p. + app. 50 p.

Ilmén M, Saloheimo A, Onnela M-L, Penttilä ME: Regulation of Cellulase Gene Expression in the Filamentous Fungus *Trichoderma reesei*. *Appl Env Microbiol.* 1997, 63:1298–1306.

Ilmén M, Thrane C, Penttilä M: The glucose repressor gene *cre1* of *Trichoderma*: Isolation and expression of a full length and truncated mutant form. *Mol Gen Genet.* 1996, 251:451–460.

Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, Yee Ho P, Kakazu Y, Sugawara K, Igarashi S, Harada S, Masuda T, Sugiyama N, Togashi T, Hasegawa M, Takai Y, Yugi K, Arakawa K, Iwata N, Toya Y, Nakayama Y, Nishioka T, Shimizu K, Mori H, Tomita M: Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* 2007, 316:593–597.

Iwatani S, Van Dien S, Shimbo K, Kubota K, Kageyama N, Iwahata D, Miyano H, Hirayama K, Usuda Y, Shimizu K, Matsui K: Determination of metabolic flux changes during fed-batch cultivation from measurements of intracellular amino acids by LC-MS/MS. *J Biotechnol*. 2007, 128:93–111.

Jamieson DJ: Oxidative stress responses of the yeast *Saccharomyces cerevisiae*. *Yeast* 1998, 14:1511–1527.

Jones JG, Bellion E: Methanol oxidation and assimilation in *Hansenula polymorpha*. An analysis by ^{13}C n.m.r. *in vivo*. *Biochem J.* 1991, 280,475–481.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y: KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 2008, 36:D480–D484.

Karp PD, Paley S, Romero P: The Pathway Tools software. *Bioinformatics* 2002, 18:S225–S232.

Kitano H: Systems biology: a brief overview. *Science* 2002, 295:1662–1664.

Kitano H: Towards a theory of biological robustness. *Mol Syst Biol*. 2007, 3:137.

Kleijn RJ, Geertman J-MA, Nfor BK, Ras C, Schipper D, Pronk JT, Heijnen JJ, van Maris AJA, van Winden WA: Metabolic flux analysis of a glycerol-overproducing *Saccharomyces cerevisiae* strain based on GC-MS, LC-MS and NMR-derived ^{13}C-labelling data. *FEMS Yeast Res.* 2007, 7:216–231.

Klipp E: Modelling dynamic processes in yeast. *Yeast* 2007, 24:943–959.

Knorr AL, Jain R, Srivastava R: Bayesian-based selection of metabolic objective functions. *Bioinformatics* 2007, 23:351–357.

Koc A, Gasch AP, Rutherford JC, Kim H-Y, Gladyshev VN: Methionine sulfoxide reductase regulation of yeast lifespan reveals reactive oxygen specis-dependent and -independent components of aging. *PNAS* 2004, 101:7999–8004.

Koch LG, Britton SL: Aerobic metabolism underlies complexity and capacity. *J Physiol.* 2008, 586:83–95.

Kolkman A, Olsthoorn MMA, Heeremans CEM, Heck AJR, Slijper M: Comparative proteome analysis of *Saccharomyces cerevisiae* grown in chemostat cultures limited for glucose or ethanol. *Mol Cell Proteomics* 2005, 4:1–11.

Kotera M, Hattori M, Oh M-A, Yamamoto R, Komeno T, Yabuzaki J, Tonomura K, Goto S, Kanehisa M: RPAIR: a reactant-pair database representing chemical changes in enzymatic reactions. *Genome Inf.* 2004, 15:P062.

Kumar R, Singh S, Singh OV: Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *J Ind Microbiol Biotechnol.* 2008, 35:377–391.

Kummel A, Panke S, Heinemann M: Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. *Molecular Systems Biology*, 2006, 2:2006.0034.

Kuorelahti S, Kalkkinen N, Penttilä M, Londesborough J, Richard P: Identification in the mold *Hypocrea jecorina* of the first fungal D-galacturonic acid reductase. *Biochemistry* 2005, 44:11234–11240.

Kwast KE, Burke PV, Poyton RO: Oxygen sensing and the transcriptional regulation of oxygen-responsive genes in yeast. *J Exp Biol.* 1998, 201:1177–1195.

Lange HC: *Quantitative Physiology of S. cerevisiae using Metabolic Network Analysis*. PhD thesis. Technical University Delft, Delft, The Netherlands; 2002.

Lange HC, Heijnen JJ: Statistical reconciliation of the elemental and molecular biomass composition of *Saccharomyces cerevisiae*. *Biotechnol Bioeng.* 2001, 75:334–44.

Lazebnik Y. Can a biologist fix a radio? Or, what I learned while studying apoptosis. *Cancer Cell* 2002, 3:179–182.

Letavayová L, Marková E, Hermanská K, Vlčková V, Vlasáková D, Chovanec M, Brozmanová J: Relative contribution of homologous recombination and nonhomologous end-joining to DNA double-strand break repair after oxidative stress in *Saccharomyces cerevisiae*. *DNA Repair* 2006, 5:602–610.

Liebermeister W, Klipp E: Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor Biol Med Model*. 2006, 3:41.

Lin S-J, Kaeberlein M, Andalis AA, Sturtz LA, Defossez P-A, Culotta VC, Fink GR, Guarente L: Calorie restriction extends *Saccharomyces cerevisiae* lifespan by increasing respiration. *Nature* 2002, 418:344–348.

Liu Z, Butow RA: Mitochondrial retrograde signalling. *Annu Rev Genet.*. 2006, 40:159–185.

López-Mirabal HR, Winther JR: Redox characteristics of the eukaryotic cytosol. *Biochim Biophysi Acta* 2008, 1783:629–640.

Maaheimo H, Fiaux J, Çakar PZ, Bailey JE, Sauer U, Szyperski T: Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional ^{13}C labelling of common amino acids. *Eur J Biochem.* 2001, 268:2464–2479.

Ma'ayan A: Insights into the organization of biochemical regulatory networks using graph theory analyses. *JBC* 2009, 284:5451–5455.

Ma H-W, Zeng A-P: The connectivity structure, giant strong component and centrality of metabolic networks. *Bioinformatics* 2003, 19:1423–1430.

Maier K, Hofmann U, Reuss M, Mauch K: Identification of metabolic fluxes in hepatic cells from transient ^{13}C-labeling experiments: Part II. Flux estimation. *Biotechnol Bioeng.* 2008 100:355–370.

Mandels M, Reese ET: Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. *J Bacteriol.* 1957, 73:269–278.

Mann M, Jensen ON: Proteomic analysis of post-translational modifications. *Nat Biotechnol*. 2003, 21:255–261.

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EGJ, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parviainen A, Schoch CL, Yao J, Barbote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS: Genome sequencing and analysis of the biomass-degrading *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol.* 2008, 26:553–560.

Minning S, Serrano A, Ferrer P, Solà C, Schmid RD, Valero F: Optimisation of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. *J Biotechnol.*,2001, 86:59–70.

Mo ML, Palsson BO, Herrgård MJ: Connecting extracellular metabolomic measurements to intracellualr flux states in yeast. *BMC Syst Biol.* 2009, 3:37.

Monod J, Wyman J, Changeux J-P: On the nature of allosteric transitions: A plausible model. *J Mol Biol.* 1965, 12:88–118.

Mu F, Williams RF, Unkefer CJ, Unkefer PJ, Faeder JR, Hlavacek WS: Carbon-fate maps for metabolic reactions. *Bioinformatics* 2007, 23:3193–3199.

Munger J, Bennett BD, Parikh A, Feng X-J, McArdle J, Rabitz HA, Shenk T, Rabinowitz: Systems-level metabolic flux profiling identifies fatty acid synthesis as a target for antiviral therapy. *Nat Biotechnol*. 2008, 26:1179–1186.

Möllney M, Wiechert W, Kownatzki D, de Graaf A: Bidirectional reaction steps in metabolic networks IV: Optimal design of isotopomer labelling systems. *Biotechnol Bioeng.* 1999, 66:86–103.

Nevoigt E: Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev*. 2008, 72:379–412.

Nielsen J, Villadsen J, Lidén G: *Bioreaction engineering principles*. Kluyver Academic, New York, 2003, p. 528.

Nikerel IE, van Winden WA, van Gulik WM, Heijnen JJ: A method for estimation of elasticities in metabolic networks using steady state and dynamic metabolomics data and linlog kinetics. *BMC Bioinformatics* 2006, 7:540.

Nissen TL, Schulze U, Nielsen J, Villadsen J: Flux distributions in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae*. *Microbiology* 1997, 143:203–218.

Notebaart RA, van Enckevort FHJ, Francke C, Siezen RJ, Teusink B: Accelerating the reconstruction of genome-scale metabolic networks. *BMC Bioinformatics* 2006, 7:296.

Nöh K, Grönke K, Luo B, Takors R, Oldiges M, Wiechert W: Metabolic flux analysis at ultra short time scale: Isotopically non-stationary ^{13}C labelling experiments. *J Biotechnol*. 2007, 129:249–267.

Oliveira GA, Tahara EB, Gombert AK, Barros MH, Kowaltowski: Increased aerobic metabolism is essential for the beneficial effects of caloric restriction on yeast life span. *J Bioenerg Biomembr.* 2008, 40:381–388.

Palmieri L, Vozza A, Agrimi G, De Marco V, Runswick J, Palmieri F, Walker JE: Identification of the Yeast Mitochondrial Transporter for Oxaloacetate and Sulfate. *JBC.*1999, 274:22184–22190.

Palsson, B. Ø. *Systems Biology: Properties of Reconstructed Networks*, 2006, Cambridge University Press, New York.

Patil KR, Nielsen J: Uncovering transcriptional regulation of metabolism by using metabolic network topology. *PNAS* 2005, 102:2685–2689.

Perrone GG, Grant CM, Dawes IW: Genetic and environmental factors influencing glutathione homeostasis in *Saccharomyces cerevisiae*. *Mol Biol Cell* 2005, 16:218–230.

Petranovic D, Nielsen J: Can yeast systems biology contribute to the understanding of human disease? *Trends Biotechnol.* 2008, 11:584-90.

Petranovic D, Vemuri GN: Impact of yeast systems biology on industrial biotechnology. *J Biotechnol.* 2009, doi:10.1016/j.jbiotec.2009.07.005, in Press.

Pfeiffer T, Schuster S, Bonhoeffer S: Cooperation and Competition in the Evolution of ATP-producing Pathways. *Science* 2001, 292:504–507.

Pharkya P, Burgard AP, Maranas CD: OptStrain: A computational framework for redesign of microbial prodcution systems. *Genome Res*. 2004, 14:2367–2376.

Pinney JW, Shirley MW, McConkey GA, Westhead DR: metaSHARK: software for automatic metabolic network prediction from DNA sequence and its application to the genomes of *Plasmodium falciparum* and *Eimeria tenella*. *Nucleic Acids Res.* 2005, 33:1399–1409.

Pitkänen E, Jouhten P, Rousu J: Inferring branching pathways in genome-scale metabolic networks with ReTrace. *BMC Syst. Biol.* 2009, 3:103.

Pitkänen E, Åkerlund A, Rantanen A, Jouhten P, Ukkonen E: ReMatch: a web-based tool to construct, store and share stoichiometric metabolic models with carbon maps for metabolic flux analysis. *J Integr Bioinform.* 2008, 5:102.

Postmus J, Canelas AB, Bouwman J, Bakker BM, van Gulik W, de Mattos MJT, Brul S, Smits GJ: Quantitative analysis of the high temperature-induced glycolytic flux increase in *Saccharomyces cerevisiae* reveals dominant metabolic regulation. *JBC* 2008, 283:23524–23532.

Price ND, Schellenberger J, Palsson BO: Uniform sampling of steady-state flux spaces: means to design experiments and to interpret enzymopathies. *Biophys J*. 2004, 87:2172–2186.

Price ND, Reed JL, Palsson BO: Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat Rev Microbiol.* 2004, 2:886–897.

Price ND, Thiele I, Palsson BO: Candidate States of *Helicobacter pylori*'s Genome-Scale Metabolic Network upon Application of "Loop Law" Thermodynamic Constraints. *Biophys J*. 2006, 90:3919–3928.

Ptacek J, Devgan G, Michaud G, Zhu H, Zhu X, Fasolo J, Guo H, Jona G, Breitkreutz A, Sopko R, McCartney RR, Schmidt MC, Rachidi N, Lee S-J, Mah AS, Meng L, Stark MJR, Stern DF, De Virgilio C, Tyers M, Andrews B, Gerstein M, Schweiter B, Predki PF, Snyder M: Global analysis of protein phosphorylation in yeast. *Nature* 2005, 438:679–684.

Qian H, Beard DA: Metabolic futile cycles and their functions: a systems analysis of energy and control. *IEE Proceedings Systems Biology* 2006, 153:192–200.

Quek L-E, Wittmann C, Nielsen LK, Krömer JO: OpenFLUX: efficient modelling software for ^{13}C-based metabolic flux analysis. *Microb Cell Fact.* 2009, 8:25.

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: A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol.* 2001, 19:45–50.

Raghevendran V, Patil KR, Olsson L, Nielsen J: Hap4 is not essential for activation of respiration at low specific growth rates in *Saccharomyces cerevisiae*. *JBC* 2006, 281:12308–12314.

Rantanen A, Mielikäinen T, Rousu J, Maaheimo H, Ukkonen E: Planning optimal measurements of isotopomer distributions for estimation of metabolic fluxes. *Bioinformatics* 2006, 22:1198–1206.

Rintala E, Toivari M, Pitkänen J-P, Wiebe MG, Ruohonen L, Penttilä M: Low oxygen levels as a trigger for enhancement of respiratory metabolism in *Saccharomyces cerevisiae*. *BMC Genomics* 2009, 10:461.

Rohde JR, Bastidas R, Puria R, Cardenas ME: Nutritional control via Tor signaling in *Saccharomyces cerevisiae*. *Curr Opin Microbiol*. 2008, 11:153–160.

Sauer U: High-Throughput phenomics: experimental methods for mapping fluxomes. *Curr Opin Biotechnol*. 2004, 15:58–63.

Savageau M: Biochemical systems analysis: III. Dynamic solutions using power-law approximation. *J Theor Biol*. 1970, 26:215–226.

Schaaff I, Heinisch J, Zimmermann FK: Overproduction of glycolytic enzymes in yeast. *Yeast* 2004, 5:285–290.

Schaub J, Mauch K, Reuss M: Metabolic flux analysis in *Escherichia coli* by integrating isotopic dynamic and isotopic stationary ^{13}C labelling data. *Biotechnol Bioeng*. 2008, 99:1170–1185.

Schellenberger J, Palsson BO: Use of randomized sampling for analysis of metabolic networks. *JBC* 2009, 284:5457–5461.

Schomburg I, Chang A, Schomburg D: BRENDA, enzyme data and metabolic information. *Nucleic Acids Res.* 2002, 30:47–49.

Schryer DW, Peterson P, Paalme T, Vendelin M: Bidirectionality and compartmentation of metabolic fluxes are revealed in the dynamics of isotopomer networks. *Int J Mol Sci.* 2009, 10:1697–1718.

Schuetz R, Kuepfer L, Sauer U: Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol Syst Biol.* 2007, 3:119.

Segrè D, Vitkup D, Church GM: Analysis of optimality in natural and perturbed metabolic networks. *PNAS* 2002, 99:15112–15117.

Selivanov VA, Marin S, Lee PWN, Cascante M: Software for dynamic analysis of tracer-based metabolomic data: estimation of metabolic fluxes and their statistical analysis. *Bioinformatics* 2006, 22:2806–2812.

Sellick CA, Reece RJ: Modulation of transcription factor function by an amino acid: activation of Put3p by proline. *EMBO J*. 2003, 22:5147–5153.

Shimizu K: Metabolic flux analysis based on ^{13}C-labeling experiments and integration of the information with gene and protein expression patterns. *Adv Biochem Eng Biotechnol.* 2004, 91:1–49.

Shinfuku Y, Sorpitiporn N, Sono M, Furusawa C, Hirasawa T, Shimizu H: Development and experimental verification of a genome-scale metabolic model for *Corynebacterium glutamicum*. *Microb Cell Fact*. 2009, 8:43.

Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, Schönfisch B, Perschil I, Chacinska A, Guiard B, Rehling P, Pfanner N, Meisinger C: The proteome of *Saccharomyces cerevisiae* mitochondria. *PNAS* 2003, 100:13207–13212.

Sola A, Maaheimo H, Ylonen K, Ferrer P, Szyperski T: Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*. *Eur J Biochem.* 2004, 271:2462–2470.

Strauss J, Mach RL, Zeilinger S, Hartler G, Stöffler G, Wolschek M, Kubicek CP: Cre1, the carbon catabolite repressor protein from *Trichoderma reesei*. *FEBS Lett.* 1995, 376:103–107.

Stephanopoulos GN, Aristidou AA, Nielsen J: *Metabolic Engineering – Principles and Methodologies*, Academic Press, San Diego 1998, 725 p.

Stryer L: *Biochemistry*, 4th edition, W.H. Freeman and Company, New York 1995, 1064 p.

Sumegi B, Sherry AD, Malloy CR, Srere PA: Evidence for orientation-conserved transfer in the TCA cycle in *Saccharomyces cerevisiae*: ^{13}C NMR studies. *Biochemistry* 1993, 32:12725–12729.

Suthers PF, Burgard AP, Dasika MS, Nowroozi F, Van Dien S, Keasling J, Maranas CD: Metabolic flux elucidation for large-scale models using ^{13}C labelled isotopes. *Metab Eng*. 2007, 9:387–405.

Swiegers J, Dippenaar N, Pretorius I, Bauer F: Carnitine-dependent metabolic activities in *Saccharomyces cerevisiae*: three carnitine acetyltranferases are essential in a carnitine-dependent strain.*Yeast* 2001, 18:585–595.

Szyperski T: Biosynthetically directed fractional 13C-labelling of proteinogenic amino acids. An efficient tool to investigate intermediary metabolism. *Eur J Biochem.* 1995, 232:433–448.

Szyperski T: ^{13}C-NMR, MS and metabolic flux balancing in biotechnology research. *Q Rev Biophys*. 1998, 31:41–106.

Szyperski T, Glaser RW, Hochuli M, Fiaux J, Sauer U, Bailey JE, Wüthrich K: Bioreaction network topology and metabolic flux ratio analysis by biosynthetic fractional ^{13}C labelling and two-dimensional NMR spectroscopy. *Metab Eng.* 1999, 1:189–197.

Tang YJ, Garcia Martin H, Myers S, Rodriquez S, Baidoo EEK, Keasling JD: Advances in analysis of microbial metabolic fluxes via (13)C isotopic labeling. *Mass Spectrom Rev*. 2009, 28:362–375.

ter Kuile BH, Westerhoff HV: Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Letters* 2001, 500:169–171.

The UniProt Consortium: The Universal Protein Resource (UniProt). *Nucleic Acids Res.* 2007, 35:D193–197.

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: Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem.* 2000, 267:5313–5329.

Toya Y, Ishii N, Hirasawa T, Naba M, Hirai K, Sugawara K, Igarashi S, Shimizu K, Tomita M, Soga T: Direct measurement of isotopomer of intracellular metabolites using capillary electrophoresis time-of-flight mass spectrometry for efficient metabolic flux analysis. *J Chromatogr* A. 2007, 1159:134–141.

Trotter E, Grant CM: Thioredoxins are required for protection against a reductive stress in the yeast *Saccharomyces cerevisiae*. *Mol Microbiol.* 2002, 46:869–878.

Uy R, Wold F: Posttranslational covalent modification of proteins. *Science* 1977, 198:890–896.

van der Brink J, Canelas AB, van Gulik WM, Pronk JT, Heijnen JJ, de Winde JH, Daran-Lapujade P: Dynamics of glycolytic regulation during adaptation of *Saccharomyces cerevisiae* to fermentative metabolism. *Appl Env Microbiol.* 2008, 74:5710–5723.

van Roermund C, Hettema E, Berg M, Tabak H, Wanders R: Molecular characterization of carnitine-dependent transport of acetyl-CoA from peroxisomes to mitochondria in *Saccharomyces cerevisiae* and identification of a plasma membrane carnitine transporter, Agp2p. *EMBO J.* 1999, 18:5843–5852.

van Winden W, Heijnen JJ, Verheijen PJT: Cumulative Bondomers: A New Concept in Flux Analysis from 2D [^{13}C,1H] COSY NMR Data. *Biotechnol Bioeng.* 2002, 80:731–745.

van Winden WA, van Dam JC, Ras C, Kleijn RJ, Vinke JL, van Gulik WM, Heijnen JJ: Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on

mass isotopomer measurements of ^{13}C-labeled primary metabolites. *FEMS Yeast Res.* 2005, 5:559–568.

van Winden WA, Wittmann C, Heinzle E, Heijnen JJ: Correcting mass isotopomer distributions for naturally occurring isotopes. *Biotechnol Bioeng.* 2002, 80:477–479.

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

Vautard G, Cotton P, Fèvre M: The glucose repressor CRE1 from *Sclerotina sclerotiorum* is functionally related to CREA from *Aspergillus nidulans* but not to the Mig proteins from *Saccharomyces cerevisiae*. *FEBS Letters* 1999, 453:54–58.

Vemuri GN, Eiteman MA, McEwen JE, Olsson L, Nielsen J: Increasing NADH oxidation reduces overflow metabolism in *Saccharomyces cerevisiae*. *PNAS* 2007, 104:2402–2407.

Verduyn C, Postma, E Scheffers WA, van Dijken JP: Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 1992, 8:501–517.

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

Visser D, Schmid J, Mauch K, Reuss M, Heijnen J: Optimal re-design of primary metabolism in *Escherichia coli* using linlog kinetics. *Metab Eng.* 2004, 6:378–390.

Visser W, van der Baan AA, Batenburg-van der Vegte W, Scheffers A, Krämer R, van Dijken JP: Involvement of mitochondria in the assimilatory metabolism of anaerobic *Saccharomyces cerevisiae* cultures. *Microbiology* 1994, 140:3039–3046.

Wahl SA, Nöh K, Wiechert W: ^{13}C labelling experiments at metabolic nonstationary conditions: An exploratory study. *BMC Bioinformatics* 2008, 9:152.

Westergaard SL, Oliveira AP, Bro C, Olsson L, Nielsen J: A systems biology approach to study glucose repression in the yeast *Saccharomyces cerevisiae*. *Biotechnol Bioeng.*2007, 96:134–145.

Weusthuis RA, Visser W, Pronk JT, Scheffers WA, van Dijken JP: Effects of oxygen limitation on sugar metabolism in yeasts – a continuous-culture study of the Kluyver effect. *Microbiology* 1994, 140:703–715.

Wiebe MG, Rintala E, Tamminen A, Simolin H, Salusjärvi L, Toivari M, Kokkonen JT, Kiuru J, Ketola RA, Jouhten P, Huuskonen A, Maaheimo H, Ruohonen L, Penttilä M: Central carbon metabolism of *Saccharomyces cerevisiae* in anaerobic,

oxygen-limited and fully aerobic steady-state conditions and following a shift to anaerobic conditions. *FEMS Yeast Res.* 2008, 8:140–154.

Wiechert W: ^{13}C metabolic flux analysis. *Metab Eng.* 2001, 3:195–206.

Wiechert W, Möllney M, Isermann N, Wurzel M, de Graaf AA: Bidirectional Reaction Steps in Metabolic Networks: III. Explicit Solution and Analysis of Isotopomer Labeling Systems. *Biotechnol Bioeng.* 1999, 66:69–85.

Wiechert W, Möllney M, Petersen S, de Graaf AA: A Universal Framework for ^{13}C Metabolic Flux Analysis. *Metab Eng.* 2001, 3:265–283.

Wild S, Roglic G, Green A, Sicree R, King H: Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 2004, 27:1047–1053.

Wittmann C: Fluxome analysis using GC-MS. *Microb Cell Fact*. 2007, 6:6.

Wittmann C, Hans M, Heinzle E: *In vivo* analysis of intracellular amino acids labelings by GC/MS. *Anal Biochem.* 2002, 307:379–382.

Woo DK, Phang TL, Trawick JD, Poyton RO: Multiple pathways of mitochondrial-nuclear communication in yeast: Intergenomic signalling involves ABF1 and affects a different set of genes than retrograde regulation. *Biochim Biophys Acta* 2009, 1789:135–145.

Zaman S, Lippman SI, Schneper L, Slonim N, Broach JR: Glucose regulates transcription in yeast through a network of signaling pathways. *Mol Syst Biol*. 2009, 5:245.

Zaman S, Lippman SI, Zhao X, Broach JR: How *Saccharomyces* responds to nutrients. *Annu Rev Genet.*. 2008, 42:2.1–2.55.

Zamboni N, Fendt S-M, Ruhl M, Sauer U: ^{13}C-based metabolic flux analysis. *Nat Protoc.* 2009, 4:878–892.

Zamboni N, Fischer E, Sauer U: FiatFlux – a software for metabolic flux analysis from ^{13}C-glucose experiments. *BMC Bioinformatics* 2005, 6:209.

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

Appendix V of this publication is not included in the PDF version. Please order the printed version to get the complete publication (http://www.vtt.fi/publications/index.jsp).

PUBLICATION I

Oxygen dependence of metabolic fluxes and energy generation of *Saccharomyces cerevisiae* CEN.PK113-1A

In: BMC Systems Biology 2008. 2:60.

BMC Systems Biology

BioMed Central

Research article

Open Access

Oxygen dependence of metabolic fluxes and energy generation of *Saccharomyces cerevisiae* CEN.PK113-1A

Paula Jouhten*, Eija Rintala, Anne Huuskonen, Anu Tamminen, Mervi Toivari, Marilyn Wiebe, Laura Ruohonen, Merja Penttilä and Hannu Maaheimo

Address: VTT Technical Research Centre of Finland, Espoo, Finland

Email: Paula Jouhten* - paula.jouhten@vtt.fi; Eija Rintala - eija.rintala@vtt.fi; Anne Huuskonen - anne.huuskonen@vtt.fi; Anu Tamminen - anu.tamminen@vtt.fi; Mervi Toivari - mervi.toivari@vtt.fi; Marilyn Wiebe - marilyn.wiebe@vtt.fi; Laura Ruohonen - laura.ruohonen@vtt.fi; Merja Penttilä - merja.penttila@vtt.fi; Hannu Maaheimo - hannu.maaheimo@vtt.fi

* Corresponding author

Published: 9 July 2008

BMC Systems Biology 2008, **2**:60 doi:10.1186/1752-0509-2-60

Received: 14 February 2008
Accepted: 9 July 2008

This article is available from: http://www.biomedcentral.com/1752-0509/2/60

Abstract

Background: The yeast *Saccharomyces cerevisiae* is able to adjust to external oxygen availability by utilizing both respirative and fermentative metabolic modes. Adjusting the metabolic mode involves alteration of the intracellular metabolic fluxes that are determined by the cell's multilevel regulatory network. Oxygen is a major determinant of the physiology of *S. cerevisiae* but understanding of the oxygen dependence of intracellular flux distributions is still scarce.

Results: Metabolic flux distributions of *S. cerevisiae* CEN.PK113-1A growing in glucose-limited chemostat cultures at a dilution rate of 0.1 h^{-1} with 20.9%, 2.8%, 1.0%, 0.5% or 0.0% O_2 in the inlet gas were quantified by ^{13}C-MFA. Metabolic flux ratios from fractional [U-^{13}C]glucose labelling experiments were used to solve the underdetermined MFA system of central carbon metabolism of *S. cerevisiae*.

While ethanol production was observed already in 2.8% oxygen, only minor differences in the flux distribution were observed, compared to fully aerobic conditions. However, in 1.0% and 0.5% oxygen the respiratory rate was severely restricted, resulting in progressively reduced fluxes through the TCA cycle and the direction of major fluxes to the fermentative pathway. A redistribution of fluxes was observed in all branching points of central carbon metabolism. Yet only when oxygen provision was reduced to 0.5%, was the biomass yield exceeded by the yields of ethanol and CO_2. Respirative ATP generation provided 59% of the ATP demand in fully aerobic conditions and still a substantial 25% in 0.5% oxygenation. An extensive redistribution of fluxes was observed in anaerobic conditions compared to all the aerobic conditions. Positive correlation between the transcriptional levels of metabolic enzymes and the corresponding fluxes in the different oxygenation conditions was found only in the respirative pathway.

Conclusion: ^{13}C-constrained MFA enabled quantitative determination of intracellular fluxes in conditions of different redox challenges without including redox cofactors in metabolite mass balances. A redistribution of fluxes was observed not only for respirative, respiro-fermentative and fermentative metabolisms, but also for cells grown with 2.8%, 1.0% and 0.5% oxygen. Although the cellular metabolism was respiro-fermentative in each of these low oxygen conditions, the actual amount of oxygen available resulted in different contributions through respirative and fermentative pathways.

Background

The yeast *Saccharomyces cerevisiae* is a facultative anaerobic organism. It is able to respond to external oxygen availability by utilizing both respirative and fermentative metabolic modes and it grows at a fast rate even when aerobic respiration is limited or completely prevented [1-3]. Metabolic response to oxygen availability requires alteration of the intracellular fluxes. The intracellular flux distribution alterations in general are mediated through transcriptional, protein level and metabolic regulation, the fluxes being the integrated network response of the regulated interactions between enzymes and metabolites [4]. Oxygen is a major determinant of the physiology of *S. cerevisiae* but understanding of the oxygen dependence of intracellular metabolic flux distributions is still scarce. Furthermore the dependence of the flux distribution on oxygen availability is of great interest in many biotechnological applications of *S. cerevisiae*, particularly those requiring a low oxygen concentration to obtain maximal product yield with simultaneous limited side products, including biomass [5].

During aerobic growth oxygen serves as a final electron acceptor in respiration. When oxygen availability is limited, cells need alternative acceptors for the electrons of NADH and $FADH_2$ to maintain the redox balance. In aerobic conditions the assimilatory NADH is oxidised mainly by the external NADH dehydrogenases or transported into mitochondria by the glycerol-3-phosphate shuttle, whereas in the absence of oxygen *S. cerevisiae* produces glycerol as a redox sink [6,7]. Since glycerol production leads to net hydrolysis of ATP and loss of carbon, *S. cerevisiae* uses oxygen preferentially for oxidation of assimilatory NADH when oxygen availability is restricted [8,9]. In addition the oxidative stress to which the cells are exposed in high external oxygen availability imposes other redox challenges.

When external oxygen availability is limited *S. cerevisiae* generates energy partially or completely through fermentation, although it is less energy efficient than respiratory metabolism [2]. While the high fermentative capacity enables *S. cerevisiae* to produce energy at a sufficient rate even in anaerobic conditions [10], constant anaerobic growth requires addition of unsaturated fatty acids and ergosterol to the culture medium since oxygen is an essential reactant in sterol biosynthesis and anabolic desaturation reactions [1,11]. Furthermore, when the respiratory system coupling NADH oxidation to the generation of a proton gradient across the mitochondrial membrane is limited, additional means for cross-membrane transport of metabolites and ions are required [1]. Growth when there is limited or no aerobic respiration thus requires an adjusment of metabolism and a major redistribution of metabolic fluxes compared to fully respiratory metabolism.

Respiration of *S. cerevisiae* becomes restricted, not only when oxygen availability is limited, but also in fully aerobic conditions when there is an excess of repressive carbon source [12-16]. The excess repressive carbon source mediates complex transcriptional regulation, including repression of respiratory genes, and thus lowers the maximal respiratory rate. Limited respiratory capacity results in alcoholic fermentation [17]. Aerobic alcoholic fermentation is also triggered at high growth rates in aerobic chemostats [18,2]. The limited respiratory capacity in both conditions has been shown to result in redistribution of intracellular carbon fluxes through respiratory and fermentative pathways [19,20,18].

Intracellular metabolic flux distributions are determined by metabolic flux analysis (MFA) which is based on stoichiometric modeling, with a system of mass balance equations for intracellular metabolites [21]. Usually the mass balance equation system is underdetermined since the number of degrees of freedom exceeds the number of measured extracellular fluxes. Linear programming can be used to solve the MFA system if a biologically meaningful objective function is formulated [22]. Including redox cofactors that are involved in all cellular metabolism into the mass balancing renders the system more constrained but requires detailed knowledge on the cofactor specificities of different isoenzymes and the relative activities of the isoenzymes in the conditions studied. This information is rarely available in the extent of a genome wide metabolic network. However, MFA with additional experimental constraints from ^{13}C-labelling experiments combined with mass spectrometry (MS) or nuclear magnetic resonance spectroscopy (NMR) detection of labelling patterns in metabolic compounds [4,18,19,23] can be used to resolve intracellular fluxes through complex pathway structures [24], including compartmentalised eukaryotic metabolic networks [25-27]. The established knowledge on the topology of the metabolic network of *S. cerevisiae* [28,29] enables modelling for MFA. The distribution of intracellular fluxes of *S. cerevisiae* to respirative and fermentative pathways in response to different reduced oxygen provisions has not been quantified with MFA combined with ^{13}C-tracer experiments before. ^{13}C-labelling has previously been used to quantify the redistribution of fluxes in *S. cerevisiae* to the respirative and fermentative pathways in response to glucose repression by comparison of batch culture fluxes to glucose-limited derepressed chemostat culture fluxes at low growth rate [19,20] and in response to high growth rates in aerobic chemostat cultures [18].

The physiology of *S. cerevisiae* in aerobic and in anaerobic conditions has been studied at the levels of gene expression [30,31], metabolite concentrations and enzyme activities [32,33], by the means of ^{13}C-metabolic flux ratio

(METAFoR) analysis [20,34], metabolic flux analysis (MFA) [9,35,36] and regulation analysis [37]. ^{13}C-tracer experiments in combination with metabolic flux analysis (MFA) have previously been applied only in studying the flux distributions of *S. cerevisiae* in aerobic glucose-limited chemostat cultures [19,38]. The effect of intermediate oxygenation conditions on *S. cerevisiae* metabolism has been the subject of some classical studies [2,3,39], including studies of the dependence of gene expression in *S. cerevisiae* on oxygenation through heme-dependent and heme-independent regulation networks, reviewed by Zitomer and Lowry (1992) and Kwast *et al.* (1998) [40,41]. Oxygen dependent transcriptional responses were observed in a range of oxygen concentrations. In addition Franzén (2003) studied ethanol production and metabolic fluxes of *S. cerevisiae* in respiratory quotient (RQ) controlled continuous cultures in a number of different microaerobic conditions by MFA without ^{13}C-tracers [9]. Franzén observed a positive correlation between biomass generation and reoxidisation of assimilatory NADH, indicating the importance of the redox balance as a determinant of the metabolic flux distribution.

The work presented here is the first where the intracellular metabolic flux distributions of *S. cerevisiae* in different levels of low external oxygen in chemostat cultures at low growth rate were quantified using ^{13}C-labelling. The low growth rate, 0.1 h^{-1}, ensured that the metabolic effects observed stemmed solely from the reduced availability of oxygen, rather than from exceeding the respiratory capacity. The flux distributions of *S. cerevisiae* central carbon metabolism under five different oxygenation conditions were solved by combining the metabolic pathway branching point constraints from the ^{13}C-labelling experiments with metabolite balancing using MFA. By including the additional constraints from the ^{13}C-labelling experiments, the cofactors could be left out from the metabolite balancing in MFA and thus the redox status regulated carbon fluxes could be reliably assessed. Completely respirative metabolism was observed in fully aerobic conditions and fully fermentative metabolism in anaerobic conditions and in the three different reduced oxygenation conditions the actual amount of oxygen available was observed to result in different flux contributions through respirative and fermentative pathways. Based on the flux distributions, energy generation of *S. cerevisiae* in the different oxygenation conditions was also determined. This paper also compares the metabolic flux distribution in different conditions of oxygen provision with the transcriptional levels of a number of metabolic genes in the same conditions, as published recently [32].

Results

S. cerevisiae CEN.PK113-1A was grown in glucose-limited chemostats at a dilution rate of 0.1 h^{-1} in five different oxygenation conditions (20.9%, 2.8%, 1.0%, 0.5% and 0.0% O_2 in the inlet gas). The corresponding average specific oxygen uptake rates (OUR) at these oxygen concentrations were 2.7, 2.5, 1.7, 1.2, and 0.0 mmol O_2 g biomass^{-1} h^{-1} as derived from a number of replicate chemostat cultivations [32]. The specific uptake rate of glucose, excretion rates of acetate, ethanol and glycerol and the biomass concentration in the different oxygenation conditions in the ^{13}C labelled replicate cultivations are given in Table 1. Net ethanol production was not observed in the aerobic cultures provided with 20.9% oxygen, indicating a fully respiratory metabolism. In 2.8% oxygen, slight ethanol excretion was observed indicating a shift to respiro-fermentative metabolism. In lower oxygen conditions, ethanol excretion rates increased further and the highest ethanol excretion rate was observed in anaerobic conditions, in which the metabolism was completely fermentative. As expected [42], the concentration of biomass was five times lower in anaerobic than in fully aerobic cultivations. Net production of glycerol was observed only in anaerobic cultivations. When only 0.5% oxygen was provided, ethanol and CO_2 yields exceeded the yield of biomass (Figure 1). In anaerobic conditions the biomass yield was only one fourth of the yield of the main product ethanol. The carbon balances closed between 96–113% in all the cultures (41.6 C-mmol/g DW [19]).

Metabolic flux ratios were determined by METAFoR analysis from the fractionally ^{13}C-labelled biomass hydrolysates by 2D NMR [34,43]. The flux ratios were calculated from the relative abundances of intact carbon backbone fragments, fragmentomers, in proteinogenic amino acids originating from a single carbon source molecule of glucose, determined from the ^{13}C-fine structures in 2D NMR spectra (Additional file 1). Flux ratios of metabolic branching points in the central carbon metabolism of *S. cerevisiae* in the different oxygenation conditions are given in Table 2. In ^{13}C-MFA metabolic flux ratios from the METAFoR analysis were used as additional constraints in a MFA system to be able to solve the metabolic net flux distribution without including the cofactors NADH and NADPH or ATP in the metabolite mass balances. The metabolic net fluxes in the different oxygenation conditions are shown in Figure 2. The confidence intervals (95%) for the net fluxes from Monte Carlo simulations of noise to the flux ratio and extracellular flux rate input data are included in Additional file 2.

Glycolytic and PPP fluxes

The METAFoR analysis showed that in fully aerobic conditions on average 32% or less of Pep originated from the PPP and the combined pool of pentose phosphates (Table 2). In lower oxygen conditions, the relative PPP flux was lower and even with 2.8% oxygen the fraction of Pep originating from pentose phosphates was only 20%. The car-

Table 1: Uptake and production rates and biomass concentration in *S. cerevisiae* CEN.PK113-1A chemostat cultures. Glucose uptake rate, glycerol, acetate and ethanol production rate and biomass concentration (mean ± SEM) of the ^{13}C-labelled glucose-limited chemostat cultures (D = 0.10 h^{-1}, pH 5.0, 30°C, 1,5 vvm gas flow) of *S. cerevisiae* CEN.PK113-1A used as input values in the ^{13}C-MFA.

	O_2 provided									
	20.9%		2.8%		1.0%		0.5%		0.0%	
	I	II	I	II	I	II	I	II	I	II
Glucose uptake rate [mmol/(g DW)$^{-1}h^{-1}$]	1.28 ± 0.04	0.87 ± 0.04	1.36 ± 0.04	1.28 ± 0.02	1.97 ± 0.02	2.13 ± 0.06	2.20 ± 0.09	2.78 ± 0.16	6.30 ± 0.25	6.58 ± 0.16
Glycerol production rate [mmol/(g DW)$^{-1}h^{-1}$]	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	1.05 ± 0.01	1.11 ± 0.03
Acetate production rate [mmol/(g DW)$^{-1}h^{-1}$]	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Ethanol production rate [mmol/(g DW)$^{-1}h^{-1}$]	0.00 ± 0.00	0.00 ± 0.00	0.07 ± 0.00	0.10 ± 0.01	1.56 ± 0.02	2.00 ± 0.08	2.59 ± 0.09	2.91 ± 0.18	9.05 ± 0.21	9.47 ± 0.31
Biomass (g DW l^{-1})	5.17 ± 0.04	5.31 ± 0.12	4.61 ± 0.09	4.86 ± 0.05	3.22 ± 0.02	2.70 ± 0.03	2.21 ± 0.03	2.08 ± 0.03	1.03 ± 0.02	0.98 ± 0.02

bon flux split ratio from G6P to glycolysis and to the oxidative branch of PPP is shown in Figure 3. The relative flux from G6P to the PPP pathway decreased as the oxygen provision was reduced. However, the results of ^{13}C MFA, shown in Figure 2, revealed that this decrease of the relative PPP flux was a result of increased glycolytic flux, while the specific flux through the oxidative branch of the PPP remained relatively constant. Progressively higher glycolytic fluxes were observed in 1.0% and in 0.5% oxygen. In anaerobic conditions the net flux in lower glycolysis remained almost the same as in 0.5% oxygen since a fraction of the carbon flux was lost in upper glycolysis to glycerol production. In anaerobic conditions the PPP flux contribution to Pep could be somewhat overestimated since the contribution of the phosphoenolpyruvate carboxykinase (PEPck) reaction to the labelling status of Pep was assumed insignificant. The flux ratio of PEPck contri-

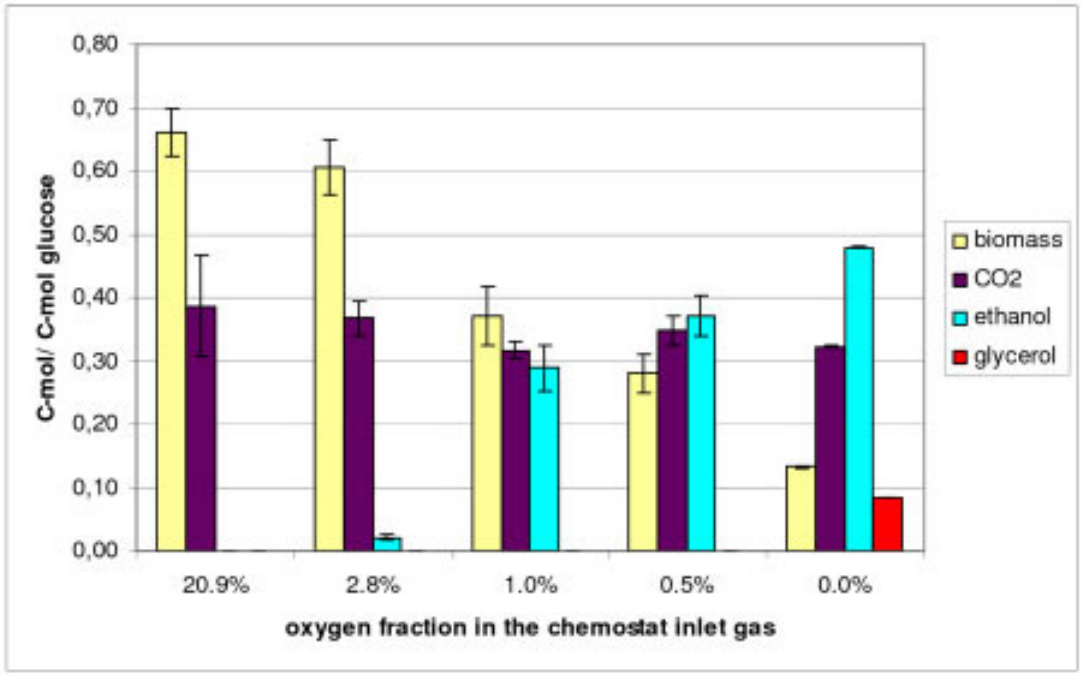

Figure 1
Average yields in *S. cerevisiae* CEN.PK113-1A glucose-limited chemostat (D = 0.1 h^{-1}) cultures. Average yields of biomass (41.6 C-mmol g biomass^{-1} [19]), CO_2, ethanol and glycerol on glucose (C-mol/C-mol) in the [U-^{13}C]glucose labelled replicate cultivations of *S. cerevisiae* CEN.PK113-1A in glucose-limited chemostat (D = 0.1 h^{-1}) in different oxygenation conditions: 20.9%, 2.8%, 1.0%, 0.5% and 0.0% oxygen of the chemostat inlet gas.

Table 2: Metabolic flux ratio (METAFoR) analysis results. Metabolic flux ratio (METAFoR) analysis results showing the origins of metabolic intermediates during growth of *S. cerevisiae* CEN.PK113-1A in glucose-limited, ^{13}C-labelled chemostat cultures (D = 0.1 h^{-1}) at different fractions of oxygen in the chemostat inlet gas. Values for two independent cultivations are given for each condition.

Metabolite	% fraction of total pool									
	20.9% O_2		2.8% O_2		1.0% O_2		0.5% O_2		0.0% O_2	
	I	II	I	II	I	II	I	II	I	II
Pep from pentose phosphates (ub)[a]	30 ± 9	34 ± 11	19 ± 6	20 ± 7	15 ± 6	19 ± 7	10 ± 7	6 ± 9	4 ± 4	4 ± 5
P5P from G3P and S7P (transketolase reaction)	51 ± 3	56 ± 6	64 ± 5	63 ± 4	82 ± 3	77 ± 3	74 ± 6	79 ± 4	86 ± 3	86 ± 5
P5P from E4P (transketolase and transaldolase)	34 ± 2	35 ± 2	27 ± 2	25 ± 2	28 ± 2	24 ± 2	26 ± 2	38 ± 2	14 ± 2	15 ± 2
Ser from Gly and C1-unit	62 ± 4	61 ± 4	61 ± 4	61 ± 4	63 ± 3	62 ± 3	62 ± 4	58 ± 3	57 ± 3	58 ± 3
Gly from CO_2 and C1-unit	4 ± 4	3 ± 3	5 ± 3	6 ± 3	4 ± 3	4 ± 3	0 ± 4	4 ± 3	4 ± 3	2 ± 3
Pep from Oaa_{cyt} (PEPck)	4 ± 7	7 ± 8	3 ± 6	1 ± 6	2 ± 10	7 ± 10	6 ± 12	0 ± 14	nd	nd
Oaa_{mit} from Pep	30 ± 2	31 ± 2	30 ± 2	29 ± 2	34 ± 2	38 ± 2	48 ± 2	57 ± 2	100 ± 2	100 ± 2
Oaa_{mit} from Oaa_{cyt}	50 ± 3	55 ± 4	52 ± 4	54 ± 3	45 ± 2	51 ± 2	60 ± 3	69 ± 2	99 ± 2	99 ± 2
Oaa_{cyt} from Pep	43 ± 2	37 ± 3	39 ± 3	35 ± 2	61 ± 3	57 ± 3	62 ± 4	59 ± 4	nd	nd
Oaa_{cyt} reversibly converted to fumarate	10 ± 7	18 ± 17	19 ± 7	17 ± 10	6 ± 5	10 ± 9	8 ± 5	14 ± 4	18 ± 7	21 ± 3
Oaa_{mit} reversibly converted to fumarate	64 ± 15	77 ± 17	71 ± 15	61 ± 14	60 ± 13	60 ± 13	62 ± 11	70 ± 8	29 ± 4	27 ± 4
Pyr_{mit} from malate (ub)[a]	3 ± 3	2 ± 6	4 ± 4	4 ± 3	0 ± 4	0 ± 4	nd[b]	nd	nd	nd
Pyr_{mit} from malate (lb)[a]	2 ± 2	1 ± 4	3 ± 3	3 ± 2	0 ± 2	0 ± 2	nd	nd	nd	nd

[a] ub upper bound, lb lower bound for the fraction of total pool, see Methods for details
[b] nd not determined, see Methods for details

bution to Pep was generally lower when oxygen provision was lower.

The METAFoR analysis also gives insight into the reversible reactions of transketolase, TK, and transaldolase, TA, (Table 2), since these reactions cleave the carbon backbone of the pentose phosphates in specific locations [43]. Higher fractions of pentose phosphates showing the reversible action of a transketolase reaction were observed when less oxygen was provided than with more. There was no clear trend in the flux through the reversible transaldolase reaction, but it was low in anaerobic cultures compared to the other conditions. The high fraction of pentose phosphates cleaved by TK and TA may reflect the proposed ping-pong mechanism of these enzymes, allowing the reaction to proceed backwards before releasing the keto- or aldogroup [44].

Anaplerosis and the Pyr branching point

The relative anaplerotic flux, the anaplerotic flux ratio defined here as the fraction of Oaa_{mit} molecules originating from Pep, was on average 31% and 30% in 20.9% and 2.8% oxygen respectively, while the relative flux from pentose phosphates to Pep was lower in 2.8% oxygen than in fully aerobic conditions (Table 2). In 1.0% oxygen the anaplerotic flux ratio was slightly higher (36%), and it was clearly higher in 0.5% oxygen (53%) than in fully aerobic conditions. Thus, in 0.5% oxygen approximately half of the Oaa_{mit} molecules originated from anaplerosis.

The carbon flux from Pyr branches into three pathways through pyruvate dehydrogenase, pyruvate decarboxylase and pyruvate carboxylase. The pyruvate dehydrogenase reaction is the first step for the carbon flux directed to the TCA cycle. The pyruvate decarboxylase reaction is the starting point for a fermentative pathway, and pyruvate carboxylase catalyses the carboxylation of Pyr to Oaa and thus also carries the anaplerotic flux. The carbon flux distributions at the pyruvate branching point were similar in fully aerobic conditions and in 2.8% oxygen (Figures 2 and 3). Slightly lower fluxes through pyruvate dehydrogenase and pyruvate carboxylase were observed in 2.8% oxygen than in 20.9% oxygen corresponding to the higher carbon flux through pyruvate decarboxylase, which reflected the production of ethanol. In conditions receiving less than 2.8% oxygen the carbon fluxes were redistributed at the pyruvate branching point and fermentative fluxes became dominating. In 1.0% oxygen the major carbon flux (on average 62%) from the pyruvate branching point was directed through pyruvate decarboxylase. The corresponding value in 0.5% oxygen was on average 77%. In 1.0% oxygen the fraction of carbon flux through pyruvate dehydrogenase (on average 23%) was less than half that observed in fully aerobic conditions (on average 65%), while the fraction of carbon flux through pyruvate carboxylase was 15% compared to 29% in fully aerobic conditions. In anaerobic conditions 94% of the carbon flux from the pyruvate branching point was directed through pyruvate decarboxylase, while pyruvate dehydrogenase flux contributed only 2% of the total flux (Figure 3).

Figure 2

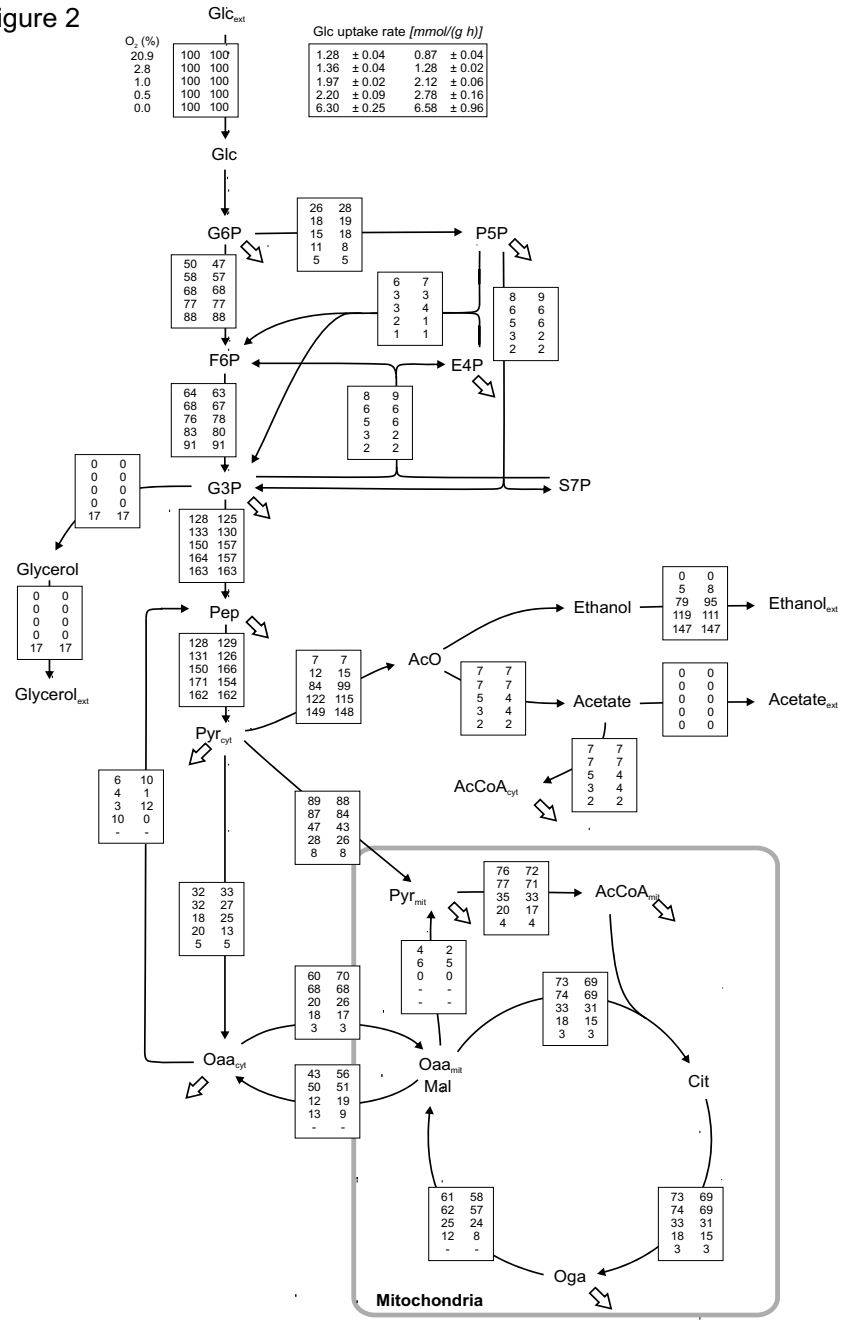

Figure 2

Metabolic net flux distributions of *S. cerevisiae* CEN.PK113-1A in different oxygenation conditions. Net flux distribution of *S. cerevisiae* CEN.PK113-1A in glucose-limited chemostat, D = 0.1 h^{-1}, in different oxygenation conditions: 20.9%, 2.8%, 1.0%, 0.5% and 0.0% oxygen of the chemostat inlet gas. The net fluxes are shown as relative fluxes normalised to the specific glucose uptake rate in the corresponding experiment. The fluxes for each reaction in the model corresponding to the 20.9%, 2.8%, 1.0%, 0.5% and 0.0% oxygen of the chemostat inlet gas are given from top to bottom and the flux values from replicate experiment are given from left to right. The specific glucose uptake rates corresponding to the different oxygenation conditions and the replicate experiments are given at the top of the figure. The net flux distributions were determined using fractional [U-^{13}C]glucose feed and metabolic flux ratio (METAFoR) analysis to obtain additional experimental constraints to render an underdetermined metabolite mass balance system solvable. The Matlab function *fmincon* performing nonlinear optimisation was applied to solve the net fluxes.

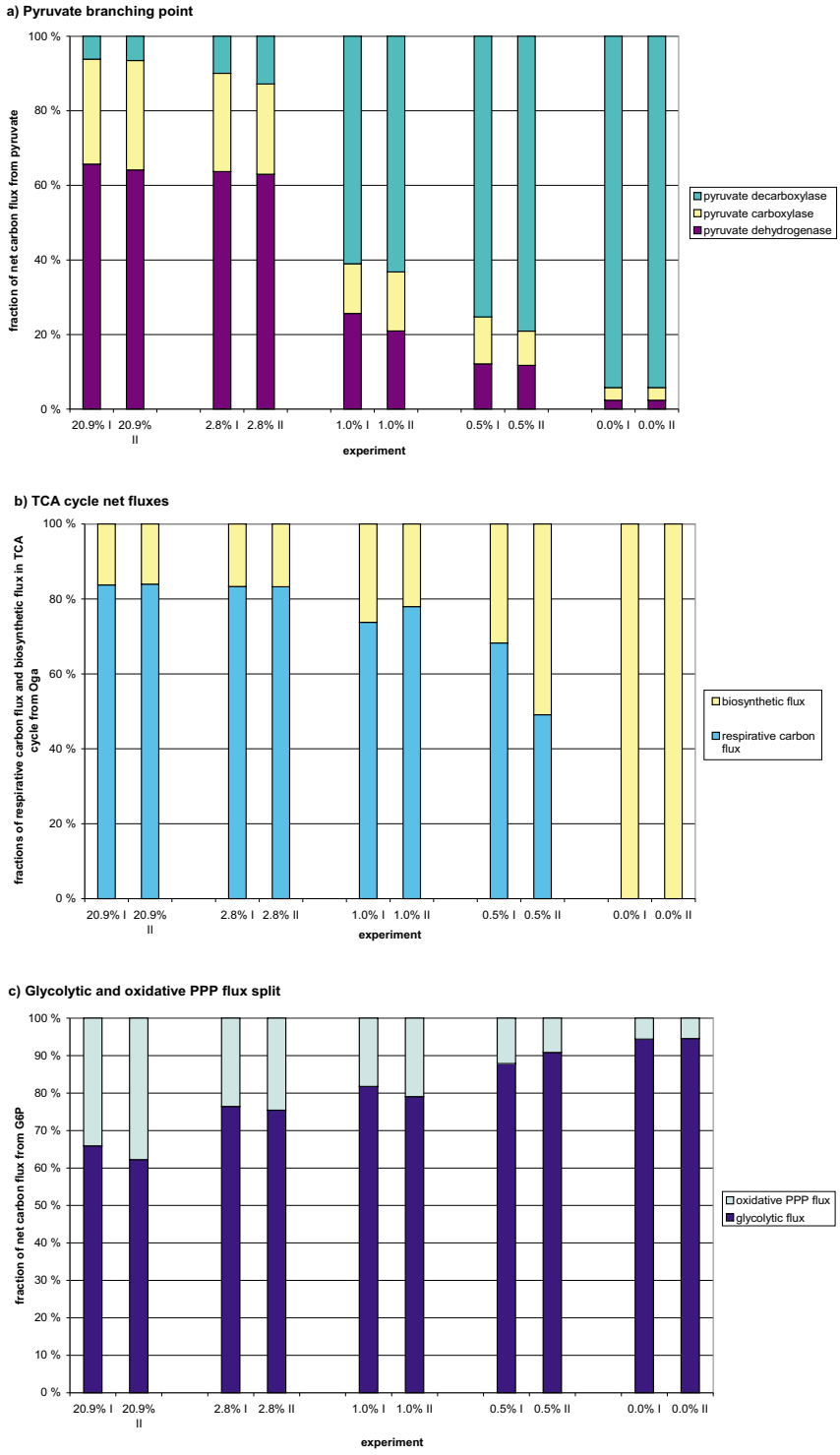

Figure 3

Fractional distributions of carbon fluxes in metabolic branching points. Fractional distribution of carbon fluxes a) from the pyruvate branching point, b) in the TCA cycle and c) from G6P to glycolysis and PPP in *S. cerevisiae* CEN.PK113-1A in glucose-limited chemostats, at D = 0.1 h^{-1}, in 20.9%, 2.8%, 1.0%, 0.5% and 0.0% oxygen of the chemostat inlet gas. Replicate experiments are indicated with I and II.

TCA cycle fluxes

Low net TCA cycle fluxes were observed in low oxygen concentrations. 2.8% oxygen in the chemostat inlet gas was enough to maintain the net TCA cycle flux at a level almost as high as in the fully aerobic conditions, but progressively lower fluxes were observed when less oxygen was provided (Figure 2). Limitation in oxygen availability reduced the respirative carbon flux through the TCA cycle, the net flux from Oga through the TCA cycle to Oaa_{mit} (x_{13}, Figure 4), whereas the specific biosynthetic flux from the TCA cycle remained constant (x_{38}, Figure 4). In fully aerobic conditions the respirative carbon flux from Oga was 84% of the net flux and even in 0.5% oxygen the respirative carbon flux was the major fraction of the net carbon flux in the TCA cycle (on average 59%) (Figure 3).

In anaerobic conditions, the C2–C3 fragments in Pep, Oaa_{cyt} and Oaa_{mit} had equal labelling patterns, as deduced from the corresponding amino acid labelling patterns (see the fragmentomer data in Additional file 1). Thus, anaplerotic flux was the only source of Oaa_{mit} (Table 2) and the TCA cycle operated as a branched pathway, with oxidative and reductive branches, instead of as a cycle [45]. The equal fractions of intact C2–C3 fragments in Pep, Oaa_{cyt} and Oaa_{mit} result in unresolved flux ratios at the metabolic branching points of Pep, Oaa_{cyt} and Pyr_{mit}. Therefore, the metabolic network model was simplified by neglecting the PEPck activity and resolving only the net transfer of Oaa across the mitochondrial membrane at the Oaa branching point. The contribution of malic enzyme flux to the Pyr_{mit} pool could not be quantified in anaerobic conditions, because the carbon fragments from the malic enzyme flux would have the same labelling pattern as the carbon fragments originating from Pyr_{cyt}. However, when 0.5% oxygen was provided, no contribution for malic enzyme flux could be observed. In anaerobic conditions, symmetrisation of Oaa as the result of reversible exchange with fumarate was observed, but this label-scrambling flux could not be quantified with the current experimental set up. The labelling pattern of Oaa_{mit} could only be partly determined from the amino acids, in which the carbon backbone originates from Oga (Table 2).

Energy metabolism

In anaerobic conditions, where no aerobic respiration is taking place, ATP is generated solely through substrate level phosphorylations. Fermentation allows redox neutral anaerobic ATP generation when acetaldehyde acts as an electron acceptor for NADH. The biosynthetic ATP requirement per biomass unit was estimated from the anaerobic ATP generating and consuming net fluxes. The specific biosynthetic ATP requirements were assumed to be constant in all oxygenation conditions, since biomass composition was assumed to be constant, as indicated in previous experimental observations which showed essentially the same biomass composition in the two extreme conditions, i.e. in fully aerobic and in anaerobic conditions [19,35]. ATP generation through substrate level phosphorylations was calculated from the reaction stoichiometry and the net flux data. The rest of the ATP demand was assumed to be provided by respirative ATP generation. The fraction of ATP generated through respiration to meet the ATP demand was 59% in fully aerobic conditions and decreased with decreasing oxygenation, as ethanol production increased (Table 3). In 0.5% oxygen, 25% of the ATP was still generated through respiration.

The efficiency of oxidative phosphorylation in different levels of oxygen was assessed by determining the P/O ratios in the different conditions. By neglecting the fraction of oxygen consumed in pathways other than respiration [46], P/O ratios were estimated from the measured OURs and the estimated amount of ATP generated through respiration (Table 3). The P/O ratios were also estimated from the electron flux to the respiratory chain and the estimated ATP generation through respiration. Assimilatory NADH generation was estimated from the anaerobic glycerol production rate, assuming that NADH generation in biosynthetic reactions was constant per g CDW^{-1} in all conditions. The generation of electron donors, NADH and $FADH_2$, in central carbon metabolism was determined from the net flux data. As NADH and $FADH_2$ are energetically equivalent in yeast, the estimated total electron flux to the respiratory chain was calculated from the summed generation of electron donors. The two estimates for P/O ratios were close to one in all the conditions.

Transcriptional regulation of metabolic enzymes

Results from Transcript analysis with the aid of Affinity Capture (TRAC) of *S. cerevisiae* in the different oxygenation conditions are presented in Wiebe *et al.* (2008) [32]. Genes encoding enzymes of central carbon metabolism were mapped to the corresponding fluxes in the metabolic network model using the *Saccharomyces* Genome Database [47] and Pearson correlation coefficients were calculated between the transcription levels of genes encoding metabolic enzymes and the fluxes through the enzymes in the different conditions. Positive correlation (> 0.60) between the transcriptional levels [32] and the corresponding fluxes in the different oxygenation conditions was found only in the respirative pathway, i.e. in pyruvate dehydrogenase and in the TCA cycle (Additional file 3).

Discussion

The dependence of the intracellular metabolic flux distribution of *S. cerevisiae* CEN.PK113-1A on the external oxygen availability was studied in glucose-limited chemostats under five different oxygenation conditions with ^{13}C-labelling. ^{13}C-labelling was utilised to obtain ratios of

Figure 4

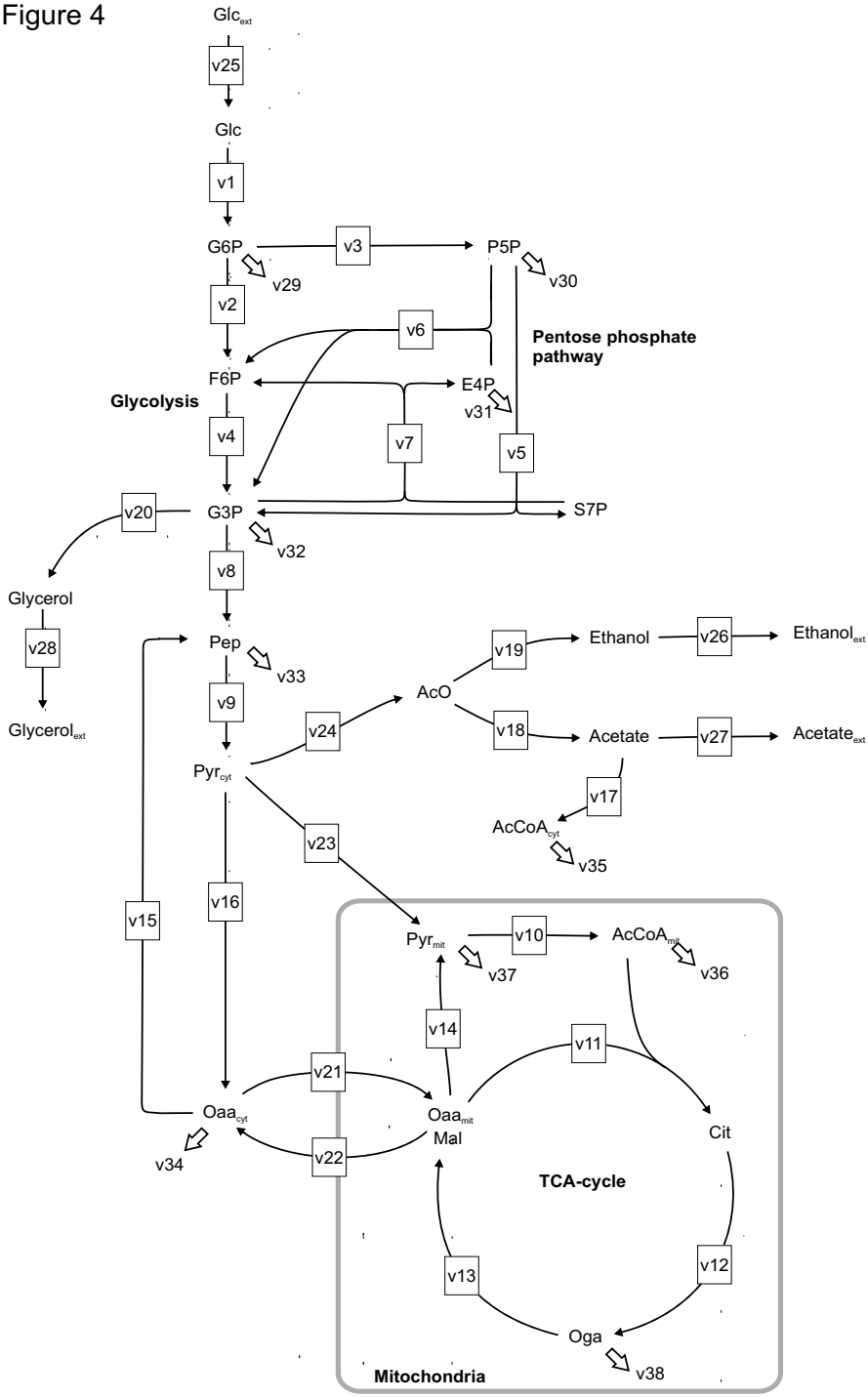

Figure 4

Metabolic network model of the central carbon metabolism of *S. cerevisiae*. Metabolic network model of the central carbon metabolism of *S. cerevisiae* was applied in the ^{13}C-metabolic flux analysis for determination of net fluxes in different oxygenation conditions. The cytosolic and mitochondrial compartments and glycolytic, pentose phosphate, TCA cycle and fermentative pathways were included in the model. The fluxes are presented as net fluxes and the directions of the arrows represent the directions of positive net fluxes. The compounds consumed or produced by external fluxes are denoted with a subscript ext. The anabolic reactions from metabolic intermediates to biosynthesis are represented by small arrows.

Table 3: Energetic features in different oxygenation conditions. Energetic features of *S. cerevisiae* CEN.PK113-1A grown in glucose-limited chemostat, D = 0.1 h-1 in different oxygenation conditions.

	O_2 provided in fermentor inlet gas				
	20.9%	2.8%	1.0%	0.5%	0.0%
ATP from respiration (%)	59	55	36	25	0
OUR (mmol $g^{-1}h^{-1}$)	2.7	2.5	1.7	1.2	0
ATP/O[a]	0.9	1.0	1.1	1.1	-
ATP/2e-[b]	1.0	0.9	1.0	1.2	-

[a]Calculated from the oxygen uptake rate (OUR).
[b]Calculated from the flux of electron donors to the respiratory chain.

intracellular fluxes at the metabolic branching points [23,43]. The flux ratio constraints were included in the MFA systems to solve the metabolic flux distributions [24]. The redox cofactors NADH and NADPH were not included in the metabolite mass balancing in ^{13}C-MFA so that the intracellular flux distributions could be reliably solved, despite the lack of precise information on the cofactor specificities and the relative activities of different isoenzymes for conditions in which redox balancing is an important determinant of cell physiology, in particular metabolic fluxes. The dilution rate in the chemostat cultivations, 0.10 h^{-1}, was well below the μ_{max} observed for the equivalent strain CEN.PK113-7D: 0.41 h^{-1} and 0.30 h^{-1} in aerobic and anaerobic conditions [15], respectively, and significantly lower than the critical dilution rate 0.27 h^{-1}, at which the metabolism of *S. cerevisiae* (CEN.PK122) has been reported to shift from fully respirative to respiro-fermentative in aerobic chemostat cultures [42]. The entirely respirative metabolism of *S. cerevisiae* under fully aerobic conditions was further confirmed by the absence of ethanol and other fermentation products in the culture supernatant and approximately the same specific rates of O_2 consumption and CO_2 production (Table 1). The controlled continuous culture conditions ensured that the metabolic effects observed under conditions of restricted respiration in the current study stemmed solely from the reduced availability of oxygen, rather than from exceeding the respiratory capacity, which has been observed to result in overflow metabolism, in aerobic alcoholic fermentation at high specific growth rate [17,18].

The switch from entirely respirative metabolism to respiro-fermentative metabolism was observed in conditions of 2.8% oxygen in the chemostat inlet gas. However, in 2.8% oxygen the respirative pathways still carried most of the carbon fluxes. When the oxygen provision was further restricted to 1.0%, thus reducing the potential of respirative ATP production, flux through the fermentative pathway increased. Since mitochondrial respiration is a significantly more efficient means to produce ATP than substrate level phosphorylations, even in only 0.5% oxygen a significant fraction (25%) of ATP was produced through respiration.

Major redistributions of carbon fluxes were observed between the different oxygenation conditions, particularly at the pyruvate branch point where the metabolism branches to three pathways. The respirative and the fermentative pathways branch out from pyruvate through the enzymes pyruvate dehydrogenase and pyruvate decarboxylase, respectively. Pyruvate decarboxylase has been found to be essential for growth on glucose in *S. cerevisiae* because of the assimilatory role of the pathway in generation of cytosolic acetyl-CoA. It is therefore also expressed during respiratory growth [48]. Wiebe *et al.* (2008) observed decreased expression of the pyruvate decarboxylase *PDC1* gene in low oxygenation [32] although the flux redistribution at the pyruvate branch point demonstrated that higher flux was directed through pyruvate decarboxylase in low than in high external oxygen, suggesting that post-transcriptional regulation is important for pyruvate decarboxylase. In 1.0% oxygen the fermentative flux through pyruvate decarboxylase became the main carbon flux from the pyruvate branch point.

Under glucose repression the respiratory pathway enzymes are severely down-regulated [17] whereas under low external oxygen availability the respiratory chain is functional but the terminal electron acceptor, oxygen, is limiting. The electron transport chain may even be optimized for low oxygen conditions by oxygen dependent modification of the terminal electron acceptor COX subunits Cox5a and Cox5b via transcriptional regulation [49]. The genes encoding TCA cycle enzymes are down-regulated in low oxygenation [32]. The carbon fluxes in the TCA cycle were also lower in lower oxygenation. In anaerobic conditions the TCA cycle operated as a branched pathway, as previously observed by Fiaux *et al.* (2003) [20] and in aerobic glucose repressed batch cultures by Gombert *et al.* (2001) [19]. On the contrary Maaheimo *et al.* (2001) [34] observed cyclic operation of the TCA cycle in aerobic batch cultures and branched operation only in anaerobic batch cultures. In mammals pyruvate dehydrogenase can be regulated via HIF1 mediated phosphorylation to reduce the flux to the TCA cycle under restricted respiration [50]. However, the activity of the *S. cerevisiae* pyruvate dehydrogenase enzyme has not been found to be regulated by phosphorylation [51].

The third flux branching from pyruvate, the anaplerotic flux, through pyruvate carboxylase, replaces the carbons lost from the TCA cycle to biosynthesis. An increase in the anaplerotic flux can be expected when the ratio of the carbon flow to biosynthesis, relative to the respirative carbon flux through the TCA cycle, is increased. When the respi-

ration rate was reduced by the reduced availability of external oxygen while the growth rate was kept constant, the respirative carbon flux was decreased whereas the specific carbon flux to biosynthesis remained the same. In 0.5% oxygen the respirative carbon flux was still over 60% of the net carbon flux to the TCA cycle whereas in anaerobic conditions there was no respirative carbon flux and the anaplerotic flux was the only source of Oaa_{mit}. Frick and Wittmann (2005) observed considerably increased anaplerotic fluxes in *S. cerevisiae* at high growth rates (D = 0.30, 0.40 h^{-1}) in aerobic chemostats compared to low growth rate (D = 0.15 h^{-1}) and the increases in the anaplerotic fluxes were accompanied by high malic enzyme fluxes [18]. High contribution of a malic enzyme flux has also been observed in aerobic glucose-repressed batch cultures [34]. In this study the highest, but still low, malic enzyme fluxes were observed in the more oxygenised conditions while the absolute anaplerotic flux remained on fairly constant level and only the ratio of anaplerosis to the TCA cycle flux was increased when oxygen concentration was reduced. Thus a high ratio of anaplerotic flux to the TCA cycle flux is associated with respiro-fermentative and anaerobic metabolism, but high absolute anaplerotic and malic enzyme fluxes with high specific growth rate and/or overflow metabolism. Overflow metabolism was not observed as a result of decreased respiratory rate achieved by reduced oxygen provision.

In fully aerobic conditions *S. cerevisiae* regenerates NAD^+ mainly through respiration. When limited oxygen availability restricts respiration, cells are forced to use other means for regeneration of $NAD^+$and mitochondrial NADH needs to be transported to the cytosol for reoxidisation. For the transport of NADH, mitochondrial alcohol dehydrogenase, encoded by *ADH3*, provides a probable redox shuttle [6,52]. *S. cerevisiae* oxidises the surplus NADH by producing glycerol as a redox sink. In this study, carbon loss to glycerol was observed only in anaerobic conditions, as expected. Based on the theoretical amount of assimilatory NADH synthesised in anaerobic conditions, 11 mmol g $biomass^{-1}$ at a growth rate 0.1 h^{-1} [10], which was consistent with the anaerobic glycerol production rate observed in this study (1.2 mmol g $biomas^{-1}$ h^{-1} [32], no net glycerol production should occur for oxygen uptake rates of 0.55 mmol O_2 g $biomass^{-1}$ h^{-1} or higher [53]. The oxygen uptake rate measured in the lowest oxygen concentration provided in this study, 0.5% O_2, was 1.2 mmol O_2 g $biomass^{-1}$ h^{-1}, which is twice the rate which would be sufficient for maintaining the cytosolic NADH balance with the external NADH dehydrogenases and mitochondrial respiration [6].

The main mechanisms in *S. cerevisiae* for mitochondrial reoxidation of cytosolic NADH are the external NADH dehydrogenases (Nde1p and Nde2p) but the glycerol-3-phosphate shuttle is also known to be active [6,7]. The anaerobic flux distribution observed was clearly different from all the other flux distributions since respiration could maintain the NADH/NAD^+ ratio in all the aerobic conditions. Weusthuis *et al.* (1994) indicated that yeasts could optimise their function for redox balancing so that available oxygen would primarily be used to maintain the redox balance, thus avoiding carbon loss to glycerol [8]. The indication has been supported by MFA modelling results of *S. cerevisiae* metabolism in low oxygen conditions [9]. Respiratory functions couple energy generation in terms of ATP with the redox balance. Since the redox cofactor NADH is one of the hub metabolites in the organism-wide network of metabolic reactions [29], the regulation of redox homeostasis encompasses all the metabolic pathways.

In this study P/O ratios were estimated in two different ways: from OURs and from the flux of electrons to the respiratory chain. The two different estimates were consistent with each other and close to one in all conditions. The effective P/O ratio has previously been estimated to be close to one in respiratory, carbon-limited cultures [54] and an increase in the effective P/O ratio in decreased respiratory fluxes has been observed in isolated mitochondria and in spheroplasts [55-57]. An ability to adjust the P/O ratio has been discussed as providing an important means to control ATP synthesis in cells to adapt to changes in energy demands [56]. In this study no significant increase in the P/O ratio was observed with decreasing respiratory fluxes.

The PPP provides precursors for biosynthesis and reductive power in the form of NADPH. The relative flux to the PPP appeared to be mainly determined by the NADPH requirement for biomass synthesis in the different oxygenation conditions. It has been stated that the flux through the PPP depends on the $NADP^+$/NADPH ratio in the cell and additionally on the $MgATP^{2-}$ pool that inhibits glucose-6-phosphate dehydrogenase, an enzyme in the oxidative branch of the PPP, allowing dynamic regulation of the relative PPP flux [58]. The dependence of the relative PPP flux on growth rate and biomass yield has been observed [18]. The relative PPP flux contributions to PEP observed with METAFoR analysis of glucose repressed cells in aerobic batch cultures [34] are essentially the same as observed in this study in response to the lowest oxygen provision and anaerobic conditions. NADPH production of the oxidative PPP in the aerobic cultivations was approximately 6 mmol g $biomass^{-1}$ in this study, assuming the maximum relative PPP flux, while approximately 9 mmol g $biomass^{-1}$ would be needed for reducing power in the form of NADPH for biomass production of yeast growing on glucose with ammonium as the nitrogen source [59]. Thus one third of the NADPH required must

have been produced in pathways other than the PPP. The isocitrate dehydrogenase reaction of the TCA cycle is assumed to be another main contributor to the production of NADPH [59]. However, NADPH is also known to be an important factor in oxidative damage prevention [60] and therefore the NADPH requirement may have been lower in the lower oxygenation conditions.

The changes in metabolic flux distribution observed in the series of different oxygenation conditions were positively correlated with the transcriptional changes of the genes encoding the flux carrying metabolic enzymes [32] only for pyruvate dehydrogenase and the TCA cycle. Glycolytic flux, in particular, showed a large increase as oxygenation was reduced, in contrast to the expression levels of some of the corresponding genes [32]. However, no extensive correlation between changes in transcription and the flux distribution in aerobic and anaerobic chemostat cultures of *S. cerevisiae* has been previously observed [61,62] and the glycolytic enzymes have recently been stated to be post-transcriptionally regulated [37,63]. In contrast, some transcriptional regulation of metabolism has been found to correlate with the glycolytic rate in batch cultures of *S. cerevisiae* strains displaying glucose uptake rates between 3.5 mmol g^{-1} h^{-1} and 15.8 mmol g^{-1} h^{-1} by Elbing *et al.* (2004) [64]. In this study, even though the specific glucose uptake rates in chemostats varied between 0.9 mmol g^{-1} h^{-1} and 6.6 mmol g^{-1} h^{-1} [32] there was no correlation with the transcriptional level of the glycolytic genes which were studied. However, in the work by Elbing *et al.* (2004), how the glycolytic rate was sensed to trigger transcriptional changes was not resolved [64]. As extensive oxygen dependent redistributions of fluxes were observed in central carbon metabolism in this work, the oxygen-dependent regulation of the fluxes in *S. cerevisiae* appears to lie mainly at the post-transcriptional level of the cell's regulatory system. However, it should be kept in mind that the oxygen dependent flux distributions of *S. cerevisiae* reflect not only the direct oxygen sensing regulatory mechanisms, but rather the ultimate response of the whole interactive multi-level regulatory system.

Conclusion

In this study the quantification of the flux distributions of *S. cerevisiae* in response to different oxygenation conditions with ^{13}C-MFA showed that the fluxes were redistributed not only between the cells grown in the fully aerobic conditions, in conditions of reduced oxygen provision and in anaerobic conditions but also for cells grown with 2.8%, 1.0% and 0.5% oxygen. Although the cellular metabolism was respiro-fermentative in each of these low oxygen conditions, the actual amount of oxygen available resulted in different contributions through respirative and fermentative pathways. The flux distribution at the pyruvate branch point, leading to respirative and fermentative pathways and to anaplerotic flux replenishing the TCA cycle, was particularly responsive to the level of reduction in oxygen provision. The oxygen-dependent regulation of the flux distribution in central carbon metabolism of *S. cerevisiae* appeared to lie mainly at the post-transcriptional level of the cell's regulatory system. Respirative pathway flux decreased progressively in reduced oxygenation conditions where the availability of terminal electron acceptor limited the respiratory rate. However, respiratory energy generation, being very efficient, provided a large fraction of the ATP required even in low oxygen conditions.

Methods

Strain and medium

Saccharomyces cerevisiae CEN.PK113-1A (*MATα*, *URA3*, *HIS3*, *LEU2*, *TRP1*, *MAL2-8c*, *SUC2*) was kindly provided by Dr. P. Kötter (Institut für Mikrobiologie, J.W. Goethe Universität Frankfurt, Germany) [65] and stored in glycerol (30% v/v) at -80°C [32].

Yeast were grown in defined minimal medium [66], with 10 g glucose l^{-1} as carbon source, and supplemented with 10 mg ergosterol l^{-1} and 420 mg Tween 80 l^{-1} (a source of oleic acid). BDH silicone antifoam (0.5 ml l^{-1}) was used to prevent foam production in the cultures [32].

Culture conditions

Cells were grown in 0.8 to 1 l medium in Biostat CT (2.5 l working volume) bioreactors. Cultures were inoculated to an initial OD_{600} of approximately 0.5 and maintained as batch cultures for 6 to 9 h, after which continuous medium feed was started while the cells were still growing exponentially. Chemostat cultures were maintained at D = 0.10 ± 0.02 h^{-1}, pH 5.0, 30°C, with 1.5 volume gas [volume culture]$^{-1}$ min^{-1} (vvm). Chemostat cultivations were performed with five different oxygenation conditions: 20.9%, 2.8%, 1.0%, 0.5% and 0.0% oxygen of the chemostat inlet gas. For cultures which received less than 20.9% O_2 in the gas stream, O_2 was replaced with the equivalent volume of N_2, so that total gas flow was maintained constant for all experiments. The N_2 gas used was 99.999% pure. Gas concentration (CO_2, O_2, N_2 and Ar) was analysed continuously in an Omnistar quadrupole mass spectrometer (Balzers AG, Liechenstein), calibrated with 3% CO_2 in Ar. ^{13}C labelled CO_2 was taken into account in the determination of CERs during feeding with ^{13}C glucose.

To achieve anaerobic conditions in the chemostat only Marprene tubing with very low oxygen permeability was used to connect the vessels. The medium reservoir was continually flushed with N_2 to prevent additional O_2 being added by diffusion into the medium in the anaerobic and low oxygen cultures. The k_La (overall oxygen transfer coefficient) for the bioreactor in the cultivation

conditions was 0.035–0.039s^{-1} (in pure water). The DOT was 83% in cultures receiving 20.9% O_2 and 0% in all the other cultures [32]. It should be noted that, based on Henry's law, the amount of oxygen able to dissolve into the medium is determined by the partial pressure of oxygen in the inlet gas, and that in oxygen-limited conditions (i.e. cultures receiving 2.8, 1.0 or 0.5% oxygen) the yeast were able to utilise all the oxygen which was able to dissolve into the medium. Oxygen was continually dissolving and continually being removed. Thus the oxygen available to the yeast was directly determined by the concentration of oxygen in the inlet gas while the measurable DOT remained zero, as also indicated by the OURs. Dissolved oxygen was measured with a Mettler Toledo InPro(R) 6000 series polarographic dissolved oxygen probe.

The culture conditions and biomass determination and chemical and metabolite analyses are described in more detail by Wiebe *et al.* (2008) [32]. The rate of ethanol loss through evaporation was estimated, based on initial measurements for 1 l cultures at 30°C, 15 vvm aeration and 800 rpm agitation and assuming that the evaporation rate would be constant in chemostat cultures under these conditions.

Biosynthetically directed fractional (BDF) ^{13}C-labelling

^{13}C-labelling experiments were performed in at least two replicate cultures under each oxygenation condition. After reaching a metabolic steady state, as determined by constant physiological parameters including biomass production, carbon dioxide evolution and oxygen uptake rates (CER and OUR), alkali utilisation, and subsequently confirmed by the observation of constant extracellular and intracellular metabolites and gene transcription, 10% of the carbon source in the medium was replaced with [U-^{13}C]glucose (Isotec, 99 atom% ^{13}C). During steady state growth the active pathways in the cells will determine how the yeast biomass becomes ^{13}C-labelled. After approximately 1.5 residence times biomass samples, 50 ml of culture broth, corresponding to 0.27 to 0.05 g CDW, were harvested by centrifugation. The cell pellets were suspended into 10 ml of 6 M HCl and the biomass was hydrolysed in sealed glass tubes at 110°C for 22 h. The suspensions were dried and dissolved in H_2O for filtration through 0.2 μm filters. The filtrates were vacuum-dried and dissolved in D_2O for NMR experiments. The pH of the samples was below 1 due to residual HCl.

As described previously [20,23,34,43,67-69], the calculation of metabolic flux ratios when using fractional ^{13}C-labelling of amino acids is based on the assumption that both a metabolic and an isotopomeric steady state exist. To establish a cost-effective protocol for a larger number of ^{13}C-labelling experiments, ^{13}C-labelled substrate was fed to a chemostat operating in a metabolic steady state for the duration of 1.5 volume changes [20,67] before harvesting the biomass. The fraction of unlabelled biomass produced prior to the start of ^{13}C-labelled medium supply was calculated following simple wash-out kinetics [69].

NMR spectroscopy

^{13}C-HSQC nuclear magnetic resonance (NMR) spectra were acquired at 40°C on a Varian Inova spectrometer operating at a ^{1}H-resonance frequency of 600 MHz essentially as described [43]. For each sample two spectra were acquired focusing on the aliphatic and aromatic regions. For the aliphatic spectra, a matrix of 1024 × 1500 (f2 × f1) complex data points was acquired and zero-filled to 4096 complex data points in f1. The spectral widths were 6000 Hz and 5100 Hz in the ^{1}H- and ^{13}C-dimensions, respectively. The narrow spectral width in the ^{13}C-dimension leads back-folding of part of the signals to the empty regions of the spectrum. For the aromatic region, a matrix of 1024 × 800 complex data points was acquired and zero-filled to 2048 complex data points in f1. The spectral widths for the aromatic spectra were 6000 Hz and 2815 Hz in the ^{1}H- and ^{13}C-dimensions, respectively. All spectra were weighted with a cosine function in both dimensions prior to the Fourier transformation. The spectra were processed using the standard Varian spectrometer software VNMR (version 6.1, C).

METAFoR analysis

Metabolic flux ratio (METAFoR) analysis was done based on the compartmentalized metabolic model of *S. cerevisiae* central carbon metabolism formulated by Maaheimo and co-workers (2001) [34]. The software FCAL (R.W. Glaser; FCAL 2.3.1) [23] was used for the integration of ^{13}C-scalar fine structures of proteinogenic amino acid carbon signals in the ^{13}C-HSQC NMR spectra and the calculation of relative abundances of intact carbon fragments originating from a single source molecule of glucose. The nomenclature used here for the intact carbon fragments, fragmentomers, has been described previously [43]. Briefly, $f^{(1)}$ represents the fraction of molecules in which the observed carbon atom and the neighbouring carbons originate from different source molecules of glucose, and $f^{(2)}$ the fraction of molecules in which the observed carbon atom and one of the two neighbouring atoms originate from the same source molecule of glucose, and $f^{(3)}$ the fraction of molecules in which the observed carbon atom and both the neighbouring carbons originate from the same source molecule of glucose. If the observed carbon exhibits significantly different ^{13}C-^{13}C scalar coupling constants with the neighbouring carbons, $f^{(2)}$and $f^{(2^*)}$ can be distinguished. The fraction of molecules with a conserved bond between the observed carbon atom and the neighbouring carbon with the smaller coupling is repre-

sented by $f^{(2)}$. $f^{(2^*)}$ then denotes for fraction of molecules where the bond is conserved between the observed carbon and the neighbouring carbon with the larger coupling. If the observed carbon is located at the end of a carbon chain, $f^{(1)}$ and $f^{(2)}$ fragmentomers can be observed indicating the conservation of the terminal two carbon fragment of the molecule.

Fragmentomer information obtained from proteinogenic amino acids can be traced back to the metabolic intermediates in central carbon metabolism since the carbon backbones of eight intermediates are conserved in amino acid synthesis pathways [34]. Mass balance equations of specific carbon fragments of the metabolic intermediates were formulated from the propagated fragmentomer information for junctions in central carbon metabolism.

Since glycolysis and the pentose phosphate pathway (PPP) are completely located in the cytosol, the upper bound for the fraction of Pep from the PPP was calculated as for prokaryotic cells and as described by Maaheimo and co-workers (2001) [34]. The fraction of Pep originating from phosphoenolpyruvate carboxykinase activity, denoted by X_{PEPck}, was calculated from the ratio of the fraction of Pep molecules containing an intact C1–C2 fragment and a cleaved bond between C2 and C3 and the fraction of Oaa_{cyt} molecules containing the equivalent fragments (Equation 1). These fragments cannot originate from glycolysis or from the PPP [34]. Phe-Cα, Tyr-Cα and Asp-Cα, Thr-Cα can be traced back to the C2 of Pep and Oaa_{cyt}, respectively, as the amino acid synthesis pathways from metabolic intermediates are known [34] (Equation 1).

$$X_{PEPck} = Pep\_10/Oaa_{cyt}\_10x = [f^{(2^*)}]\{phe, Tyr\text{-}C\alpha\}/[f^{(2^*)}]\{Asp, Thr\text{-}C\alpha\} \quad (1)$$

where *Pep*_10 denotes for the fraction of Pep molecules that possess an intact C1–C2 bond and cleaved C2–C3 bond and Oaa_{cyt}_10*x* denotes for the fraction of Oaa_{cyt} molecules that possess an intact C1–C2 bond, a cleaved C2–C3 bond and either intact or cleaved C3–C4 bond.

The Oaa_{mit} molecules originating from Oga through the TCA cycle possess cleaved C2–C3 bonds. The fraction of Oaa_{mit} originating from transport over the mitochondrial membrane from Oaa_{cyt} was solved from a mass balance of intact C2–C3 fragments in Oaa_{mit}. The conserved connectivity of the C2–C3 fragment in Oaa_{mit} can be propagated back from Glu-Cα and Pro-Cα carbons that represent the C2 carbon in Oga, since the C2–C3 fragment of Oaa_{mit} is conserved in the TCA cycle as the C2–C3 fragment of Oga. The fraction of Oaa_{mit} from Oaa_{cyt}, denoted by $X_{Oaa\text{-}transport}$, was calculated as a ratio of intact C2–C3 fragments in Oga and Oaa_{cyt} (Equation 2).

$$X_{Oaa-transport} = Oga\_x1xx / Oaa_{cyt}\_x1x = \left(\left[f^{(2)} + f^{(3)}\right]\{Glu, \mathrm{Pro} - C\alpha\}\right) / \left(\left[f^{(2)} + f^{(3)}\right]\{Asp, Thr - C\alpha, Asp - C\beta\}\right) \quad (2)$$

The fraction of Oaa_{cyt} originating from Pyr_{cyt}, denoted by $X_{Oaa_{cyt}\_from\_Pyr_{cyt}}$, was solved from the mass balance of intact C2–C3 fragments (Equation 3). Since the flux from Pep to Pyr_{cyt} through phosphoenolpyruvate kinase and further through pyruvate carboxylase to Oaa_{cyt} can be assumed to be irreversible in the experimental conditions used here, the C2–C3 fragments of Pep were used in the mass balance equations. The conserved connectivity of the C2–C3 fragment in Pyr_{cyt} can be observed from Phe-Cα and Tyr-Cα that represent the C2 carbon of Pep (Equation 3).

$$X_{Oaa_{cyt}\_from\_Pyr_{cyt}} = \left(Oaa_{cyt}\_x1x - Oga\_x1xx\right) / \left(Pep\_x1 - Oga\_x1xx\right) = \frac{\left[f^{(2)}+f^{(3)}\right]\{Asp, Thr-C\alpha, Asp-C\beta\} - \left[f^{(2)}+f^{(3)}\right]\{Glu, \mathrm{Pro}-C\alpha\}}{\left[f^{(2)}+f^{(3)}\right]\{Phe, Tyr-C\alpha, C\beta\} - \left[f^{(2)}+f^{(3)}\right]\{Glu, \mathrm{Pro}-C\alpha\}} \quad (3)$$

The upper and lower bounds for Pyr_{mit} originating from the malic enzyme reaction, denoted by X_{MAE_ub} and X_{MAE_lb} respectively, were calculated from a mass balance of intact C2–C3 fragments of Pyr_{mit} (Equations 4 and 5). The upper and lower bounds were obtained from the assumption that the substrate fragment for malic enzyme has an equally conserved connectivity as Oga and Oaa_{mit}. The intact fragments in Oaa_{mit} were obtained from the intact fragments in Oga since the C2–C3–C4 fragment of Oaa_{mit} is conserved in the TCA cycle in synthesis of Oga. The intact fragments in biosynthetic precursor Oga were deduced from the *f*-values of Glu and Pro carbons (Equations 4 and 5).

$$X_{MAE\_ub} = \left(Pep\_x1 - Pyr_{mit}\_x1\right) / \left(Pep\_x1 - Oga\_x1xx\right) = \frac{\left[f^{(2)}+f^{(3)}\right]\{Phe, Tyr-C\alpha, C\beta\} - \left[f^{(2)}+f^{(3)}\right]\{Ala-C\alpha, C\beta\}}{\left[f^{(2)}+f^{(3)}\right]\{Phe, Tyr-C\alpha, C\beta\} - \left[f^{(2)}+f^{(3)}\right]\{Glu, \mathrm{Pro}-C\alpha\}} \quad (4)$$

$$X_{MAE\_lb} = \left(Pep\_x1 - Pyr_{mit}\_x1\right) / Pep\_x1 = \frac{\left[f^{(2)}+f^{(3)}\right]\{Phe, Tyr-C\alpha, C\beta\} - \left[f^{(2)}+f^{(3)}\right]\{Ala-C\alpha, C\beta\}}{\left[f^{(2)}+f^{(3)}\right]\{Phe, Tyr-C\alpha, C\beta\}} \quad (5)$$

^{13}C-MFA

Metabolic flux analysis (MFA) was used to determine intracellular metabolic fluxes, with METAFoR analysis providing additional experimental constraints to solve the MFA system [24]. A stoichiometric model of central carbon metabolism of *S. cerevisiae* was formulated, based on the model used in the METAFoR analysis [34] (Additional file 4). The model included the glycolytic and the pentose phosphate pathways, the TCA cycle and the fermentative pathways, production of glycerol and anabolic fluxes from metabolic intermediates to biosynthesis. The glyoxylate cycle was omitted from the model since the METAFoR data showed that the pathway was inactive. The labelling pattern of succinate that would have originated from the glyoxylate cycle was calculated from Asp and Lys fragmentomers representing the labelling patterns of Oaa_{cyt} and $AcCoA_{cyt}$ respectively. No trace of influx of succinate originating from the Glyoxylate cycle to the TCA cycle was observed since the labelling pattern of Oga derived from Glu fragmentomers was fully explained by the TCA carbon flux. Separate pools of Pyr, AcCoA and Oaa in the two cellular compartments, cytoplasm and mitochondria, were included in the model. Mal was lumped in the same pool with Oaa_{mit}. Also the pentose phosphates formed a single pool and the triose phosphates were combined in the pools of G3P and Pep. DHAP, the precursor for glycerol synthesis, was also combined with the G3P pool. TCA cycle metabolites were represented by the pools of citrate, Oga and Oaa_{mit}. Scrambling of ^{13}C-labels in the symmetric molecules succinate and fumarate was taken into account. The transport of Pyr and Oaa across the mitochondrial membrane were included in the model but the transport of AcCoA, the final step of the cytosolic PDH bypass, was omitted since exogenous carnitine would be required for the carnitine shuttle to be active [70-72], and carnitine was not provided in the medium. In addition carnitine acetyltransferase activity has not been detected in *S. cerevisiae* grown in anaerobic chemostats at 0.1 h^{-1} [35]. Since acetaldehyde can freely diffuse across the mitochondrial membrane and acetaldehyde dehydrogenase (EC 1.2.1.3) and AcCoA synthetase (EC 6.2.1.1) enzymes have both been isolated in the mitochodrial proteome [73], PDH bypass could also be partially located in mitochondria and contribute directly to the formation of $AcCoA_{mit}$. In absence of fluxes inducing significantly dissimilar labelling patterns to the C2–C3 fragments of Pyr_{cyt} and Pyr_{mit} i.e. in conditions of low malic enzyme fluxes as observed in this study, ^{13}C-labelling cannot solely reveal the possible contribution of PDH bypass pathways to the carbon flux to mitochondria. However, in the cultivations performed, the expression of *ACS1* encoding the mitochondrial AcCoA synthetase, essential for the contribution of mitochondrial PDH bypass to the formation of $AcCoA_{mit}$, was negligible and the expression of *ACS2* encoding the cytosolic isoenzyme was substantially higher [32]. Thus, the mitochondrial PDH bypass was not included in the model.

The metabolic fluxes were modelled as net fluxes so that a net flux in the forward direction was assigned with a positive value and a net flux in the reverse direction was assigned with a negative value. As an exception, the transport of Oaa across the mitochondrial membrane was modelled as two one-directional transport reactions. In *S. cerevisiae* the transport of OAA across the mitochondrial membrane can occur via mitochondrial Oaa transporter OAC1 facilitated transport [74].

The stoichiometric model for experiments in 20.9%, 2.8% and 1.0% oxygen conditions consisted of 38 reactions coupling 34 metabolites including duplicated extracellular metabolites and uptake and production fluxes, Figure 4. The 14 fluxes across the system boundary included glucose uptake and excretion fluxes of ethanol, acetate and glycerol and the fluxes of the metabolic precursors to macromolecule synthesis for biomass production. The METAFoR results were used to identify inactive reactions, to constrain the stoichiometric models for the experiments with 0.5% and 0.0% oxygen by omitting inactive fluxes. The stoichiometric model for experiments in 0.5% oxygen consisted of 37 reactions, coupling 34 metabolites and excluding the malic enzyme activity from the first model of the network of active reactions. The compartmentalization of central carbon metabolism in anaerobic conditions is evident from the vital anabolic role of mitochondria in the absence of oxygen [75]. However, in completely anaerobic conditions only the net transport of Oaa across the mitochondrial membrane is resolvable and the activities of PEPck and malic enzyme reactions cannot be quantified. Since, according to the METAFoR analysis, the PEPck reaction showed only slight activity in the other conditions studied and its activity decreased as the oxygen provided was reduced, it was omitted from the anaerobic stoichiometric model. *MAE1* has been shown to be induced in anaerobic conditions and its possible role in provision of NADPH in mitochondria in anaerobic conditions has been discussed [76]. However, the malic enzyme reaction also showed only slight activity in all the conditions where quantification was possible and had its lowest activity in 0.5% oxygen. Thus, the malic enzyme reaction was omitted from the anaerobic model. Under anaerobic conditions the stoichiometric model of the active pathways consisted of 34 reactions and 34 metabolites.

After including the measured uptake and excretion rates and the rates of metabolic precursor depletion to biomass synthesis, as determined from the composition of *S. cerevisiae* biomass previously reported [19], in the models, the linear equation systems remained underdetermined. The

composition of *S. cerevisiae* biomass was assumed to be the same in all the conditions studied, since the biomass composition in the two extreme conditions, i.e. in fully aerobic and in anaerobic conditions, has been experimentally shown to be essentially the same [19,35]. Solvable systems were obtained by further constraining the MFA systems with one to six linearly independent constraints, depending on the structure of the network of active reactions from the METAFoR analysis as described by Fischer and co-workers (2004) [24]. Using the constraints from the METAFoR analysis, it was not necessary to include redox cofactor mass balances in the mass balance constraints in ^{13}C MFA. Cofactor mass balances are sources of errors since the correct balancing requires detailed knowledge of the relative activities of different isoenzymes and the enzyme cofactor specificities on a cell wide scale. The mass balances of the metabolites were formulated as a linear equation system as described in [24] (Equation 6):

$$N_i x - b = R_m \quad (6)$$

where N_i is the stoichiometric matrix of the active network i determined from the METAFoR analysis fragmentomer data, x is the flux distribution vector, b is the vector of the measured extracellular fluxes and R_m is the vector of the residuals of metabolite mass balances.

The flux ratio equations were set up according to the METAFoR analysis for the reactions in the stoichiometric models of the central carbon metabolism of *S. cerevisiae* (Equations 7 to 11, the reaction numbers are defined in Figure 4). Depending on the structure of the network of active reactions the flux ratio equations included one to six of the following (Equations 7 to 11):

the fraction of Pep from PPP assuming a maximal contribution of PPP

$$fr1 = \frac{x_5+3x_6+2x_7}{x_5+2x_4+x_6} \quad (7)$$

the fraction of Pep originating from Oaa_{cyt}, X_{PEPck}:

$$fr2 = \frac{x_{15}}{x_{15}+x_8} \quad (8)$$

the fraction of Oaa_{mit} originating from Oaa_{cyt}, $X_{Oaa\text{-}transport}$:

$$fr3 = \frac{x_{21}}{x_{21}+x_{13}} \quad (9)$$

the fraction of Oaa_{cyt} originating from Pyr_{cyt}, $X_{Oaa_{cyt}\_from\_Pyr_{cyt}}$:

$$fr4 = \frac{x_{16}}{x_{16}+x_{22}} \quad (10)$$

the upper and lower bounds for Pyr_{mit} originating from the malic enzyme reaction, X_{MAE_ub} and X_{MAE_lb}:

$$fr5 \le \frac{x_{14}}{x_{14}+x_9} \le fr6 \quad (11)$$

The following linear constraint equations were obtained from the flux ratio equations and included to the MFA systems to solve the metabolite mass balances (Equations 12 to 17):

$$x_5 + 3x_6 + 2x_7 - fr1(x_5 + 2x_4 + x_6) = 0 \quad (12)$$

$$x_{15} - fr\,2(x_{15} + x_8) = 0 \quad (13)$$

$$x_{21} - fr3(x_{21} + x_{13}) = 0 \quad (14)$$

$$x_{16} - fr4(x_{16} + x_{22}) = 0 \quad (15)$$

$$x_{14} - fr5(x_{14} + x_9) = 0 \quad (16)$$

$$fr6(x_{14} + x_9) - x_{14} = 0 \quad (17)$$

Irreversibility was assumed for the intracellular fluxes x_3, x_4, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{21}, x_{22}, x_{23}, x_{24}, for extracellular fluxes x_{25}, x_{26}, x_{27}, x_{28}, and for the depletion of precursors to biosynthetic reactions and thus, only positive values were allowed for the fluxes. The minimization of the sum of the weighted square residuals of the metabolite mass balances was done using the Matlab function *fmincon*. The extracellular metabolite mass balances were assigned weights according to the experimental measurement error estimates. The biomass precursor metabolite mass balances were assigned ten-fold larger weights, relative to their stoichiometric coefficients in the biomass composition, since the biomass composition was assumed constant in all the conditions studied [2]. The flux ratio constraints were included as strict constraints. The optimization was started 10000 times from random initial values to evaluate the uniqueness of the optimal flux distribution. The sensitivity of the flux distribution solutions to the noise in the flux ratio data and in the extracellular flux data was studied by Monte Carlo-simulations [77]. Additive normally distributed noise with zero mean and experimentally determined variances of the flux ratios and the extracellular fluxes was sampled to the flux ratios and to the extracellular flux data, separately and simultaneously. A flux distribution was solved for each of the 100 sets of input data from 12 random initial

flux distributions. Confidence intervals (95%) for the fluxes were determined.

Abbreviations

AcCoA: acetyl coenzyme A; AcO: acetaldehyde; CDW: cell dry weight; Cit: citrate; DHAP: dihydroxyacetone phosphate; E4P: erythrose 4-phosphate; F6P: fructose 6-phosphate; G3P: glyceraldehyde 3-phosphate; G6P: glucose 6-phosphate; HSQC: heteronuclear single quantum coherence; Mal: malate; METAFoR: metabolic flux ratio; Oaa: oxaloacetate; Oaa_{cyt}: cytosolic oxaloacetate; $Oaa_{mit:}$ mitochondrial oxaloacetate; Oga: 2-oxoglutaric acid; Pep: phosphoenolpyruvate; PEPck: phosphoenolpyruvate carboxykinase; PPP: pentose phosphate pathway; Pyr: pyruvate; $Pyr_{cyt:}$ cytosolic pyruvate; $Pyr_{mit:}$ mitochondrial pyruvate; P5P: pentose 5-phosphate; S7P: sedoheptulose 7-phosphate; TCA: tricarboxylic acid.

Authors' contributions

PJ participated in the design of the study, performed chemostat cultivations, carried out the NMR experiments, designed and performed the modeling and the computational work and drafted the manuscript, ER and MT participated in the design of the study and performed chemostat cultivations, AH and AT performed chemostat cultivations, MW participated in the design of the study, performed chemostat cultivations and the calculations of the physiological parameters and revised the language of the manuscript, LR and MP participated in the design of the study, HM participated in the design of the study and the NMR experiments and helped to draft the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1

***Relative abundances of intact carbon fragments in proteinogenic amino acids**. Relative abundances of intact C2 and C3 fragments (f-values) in proteinogenic amino acids describing the conservation of carbon chain fragments in the metabolism of* S. cerevisiae *CEN.PK113-1A in glucose-limited chemostat, D = 0.1 h^{-1}, in different oxygenation conditions. The fragmentomers were obtained using biosynthetic fractional [U-^{13}C]glucose labelling during metabolic steady state, ^{13}C-HSQC NMR measurements and software FCAL for the integration of ^{13}C-scalar fine structures of the amino acid carbon signals in the ^{13}C-HSQC NMR spectra and the calculation of relative abundances of fragmentomers using probabilistic equations relating the ^{13}C-scalar fine structures and the intact carbon chain fragments [43]. The* f-values *for the replicate experiments and their standard deviations are given in columns. For nomenclature of* f-values, *see Methods.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1752-0509-2-60-S1.doc]

Additional file 2

***Confidence intervals (95%) for the metabolic net fluxes**. 95% confidence intervals for the net fluxes in the central carbon of* S. cerevisiae *CEN.PK113-1A in glucose-limited chemostats, D = 0.1 h^{-1}, in different oxygenation conditions. The lower and upper bounds for the confidence intervals of the replicate experiments are given in µmol/(g CDW h) for each net flux. ND stands for not determined because the reactions were excluded from the stoichiometric models according to the metabolic flux ratio (METAFoR) analysis data.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1752-0509-2-60-S2.doc]

Additional file 3

***Pearson correlations between the transcriptional levels of genes encoding metabolic enzymes and the corresponding fluxes**. The Pearson correlation between the transcriptional levels of the* S. cerevisiae *genes encoding metabolic enzymes, whose levels were measured in Wiebe* et al. *[32], mapped to the corresponding fluxes in the ^{13}C-MFA model, and flux values in five different oxygenation conditions (0, 0.5, 1.0, 2.8 and 20.9% oxygen). The pairs of genes and fluxes that showed positively correlated behaviour (Pearson correlation >0.60) are marked with green.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1752-0509-2-60-S3.pdf]

Additional file 4

***Stoichiometric model of the central carbon metabolism of* S. cerevisiae**. *Reactions in the stoichiometric model of the central carbon metabolism of* Saccharomyces cerevisiae, *including also anabolic fluxes from metabolic intermediates to biosynthesis, transport reactions across the mitochondrial membrane and uptake and excretion reactions, applied in the ^{13}C-MFA determination of the metabolic net fluxes in different oxygenation conditions.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1752-0509-2-60-S4.doc]

Acknowledgements

This work was supported by the Academy of Finland (Centre of Excellence, Industrial Biotechnology 2000–2005; projects number 214568 and 204369, and SYSBIO programme; project number 207435) and Tekes, the Finnish Funding Agency for Technology and Innovation (Project numbers 40135/04 and 40537/05).

We thank Outi Könönen, Pirjo Tähtinen, Eila Leino and Tarja Hakkarainen for technical assistance. We acknowledge Ph.D. Ari Rantanen for advice and discussions on metabolic flux analysis.

References

1. Visser W, Scheffers A, Batenburg-van der Vegte WH, van Dijken J: **Oxygen Requirements of Yeasts.** *Appl Environ Microbiol* 1990, **56(12):**3785-3792.
2. Furukawa K, Heinzle E, Dunn IJ: **Influence of Oxygen on the Growth of *Saccharomyces cerevisiae* in Continuous Culture.** *Biotechnol Bioeng* 1983, **25:**2293-2317.
3. Oura E: *The effect of aeration on the growth energetics and biochemical composition of baker's yeast* PhD thesis. Helsinki University, Helsinki, Finland; 1972.
4. Sauer U: **Metabolic networks in motion: ^{13}C-based flux analysis.** *Mol Syst Biol* 2006, **2:**62.

5. van Maris AJA, Konings WN, van Dijken JP, Pronk JT: **Microbial export of lactic and 3-hydroxypropanoic acid: implications for inductrial fermentation processes.** *Metab Eng* 2004, **6:**245-255.
6. Bakker BM, Overkamp KM, van Maris AJA, Kötter P, Luttik MAH, van Dijken JP, Pronk JT: **Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*.** *FEMS Microbiol Rev* 2001, **25:**15-37.
7. Rigoulet M, Aguilaniu H, Avéret N, Bunoust O, Camougrand N, Grandier-Vazeille X, Larsson C, Pahlman IL, Manon S, Gustafsson L: **Organization and regulation of the cytosolic NADH metabolism in the yeast *Saccharomyces cerevisiae*.** *Mol Cell Biochem* 2004, **256-257:**73-81.
8. Weusthuis RA, Visser W, Pronk JT, Scheffers WA, van Dijken JP: **Effects of oxygen limitation on sugar metabolism in yeasts – a continuous-culture study of the Kluyver effect.** *Microbiology* 1994, **140:**703-715.
9. Franzèn CJ: **Metabolic flux analysis of RQ-controlled microaerobic ethanol production by *Saccharomyces cerevisiae*.** *Yeast* 2003, **20:**117-132.
10. Verduyn C, Postma E, Scheffers WA, van Dijken JP: **Physiology of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures.** *J Gen Microbiol* 1990, **136:**395-403.
11. Snoek I, Steensma Y: **Factors involved in anaerobic growth of *Saccharomyces cerevisiae*.** *Yeast* 2007, **24:**1-10.
12. Gancedo JM: **Yeast carbon catabolite repression.** *Microbiol Mol Biol Rev* 1998, **62:**334-361.
13. De Deken RH: **The Crabtree effect: a regulatory system in yeast.** *J Gen Microbiol* 1966, **44:**149-156.
14. Rieger M, Käppeli O, Fiechter A: **The Role of Limited Respiration in the Incomplete Oxidation of Glucose by *Saccharomyces cerevisiae*.** *J Gen Microbiol* 1983, **129:**653-661.
15. van Hoek P, van Dijken JP, Pronk JT: **Regulation of fermentative capacity and levels of glycolytic enzymes in chemostat cultures of *Saccharomyces cerevisiae*.** *Enzyme Microb Technol* 2000, **26:**724-736.
16. Westergaard SL, Oliveira AP, Bro C, Olsson L, Nielsen J: **A systems biology approach to study glucose repression in the yeast *Saccharomyces cerevisiae*.** *Biotechnol Bioeng* 2007, **96:**134-145.
17. Vemuri GN, Eiteman MA, McEwen JE, Olsson L, Nielsen J: **Increasing NADH oxidation reduces overflow metabolism in *Saccharomyces cerevisiae*.** *PNAS* 2007, **104:**2402-2407.
18. Frick O, Wittmann C: **Characterization of the metabolic shift between oxidative and fermentative growth in *Saccharomyces cerevisiae* by comparative ^{13}C flux analysis.** *Microb Cell Fact* 2005, **4:**30-46.
19. Gombert AK, dos Santos MM, Christensen B, Nielsen J: **Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression.** *J Bacteriol* 2001, **183:**1441-1451.
20. Fiaux J, Ìakar PZ, Sonderegger M, Wüthrich K, Szyperski T, Sauer U: **Metabolic-Flux Profiling of the Yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*.** *Eukaryot Cell* 2003, **2:**170-180.
21. Stephanopoulos GN, Aristidou AA, Nielsen J: *Metabolic Engineering: Principles and Methodologies* Academic Press, San Diego, California; 1998.
22. Schuetz R, Kuepfer L, Sauer U: **Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*.** *Mol Syst Biol* 2007, **3:**119.
23. Szyperski T, Glaser RW, Hochuli M, Fiaux J, Sauer U, Bailey JE, Wüthrich K: **Bioreaction network topology and metabolic flux ratio analysis by biosynthetic fractional ^{13}C labelling and two-dimensional NMR spectroscopy.** *Metab Eng* 1999, **1:**189-197.
24. Fischer E, Zamboni N, Sauer U: **High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived ^{13}C constraints.** *Anal Biochem* 2004, **325:**308-316.
25. Blank LM, Kuepfer L, Sauer U: **Large-scale ^{13}C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast.** *Genome Biol* 2005, **6:**R49.
26. Blank LM, Lehmbeck F, Sauer U: **Metabolic-flux and network analysis of fourteen hemiascomycetous yeasts.** *FEMS Yeast Res* 2005, **5:**545-558.
27. Fredlund E, Blank LM, Schnürer J, Sauer U, Passoth V: **Oxygen- and glucose-dependent regulation of central carbon metabolism in *Pichia anomala*.** *Appl Environ Microbiol* 2004, **70:**5905-5911.
28. Duarte NC, Herrgård MJ, Palsson BO: **Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model.** *Genome Res* 2004, **14:**1298-1309.
29. Förster J, Famili I, Fu P, Palsson BO, Nielsen J: **Genome-Scale Reconstruction of the *Saccharomyces cerevisiae* Metabolic Network.** *Genome Res* 2003, **13:**244-253.
30. Tai SL, Boer VM, Daran-Lapujade P, Walsh MC, de Winde JH, Daran JM, Pronk JT: **Two-dimensional Transcriptome Analysis in Chemostat Cultures.** *J Biol Chem* 2005, **280:**437-447.
31. Ter Linde JJM, Liang H, Davis RW, Steensma HY, van Dijken JP, Pronk JT: **Genome-Wide Transcriptional Analysis of Aerobic and Anaerobic Chemostat Cultures of *Saccharomyces cerevisiae*.** *J Bacteriol* 1999, **181:**7409-7413.
32. Wiebe MG, Rintala E, Tamminen A, Simolin H, Salusjärvi L, Toivari M, Kokkonen JT, Kiuru J, Ketola RA, Jouhten P, Huuskonen A, Maaheimo H, Ruohonen L, Penttilä M: **Central carbon metabolism of *Saccharomyces cerevisiae* in anaerobic, oxygen-limited and fully aerobic steady-state conditions and following a shift to anaerobic conditions.** *FEMS Yeast Res* 2008, **8:**140-154.
33. Smits HP, Hauf J, Muller S, Hobley TJ, Zimmermann FK, Hahn-Hägerdal B, Nielsen J, Olsson L: **Simultaneous overexpression of enzymes of the lower part of glycolysis can enhance the fermentative capacity of *Saccharomyces cerevisiae*.** *Yeast* 2000, **16:**1325-1334.
34. Maaheimo H, Fiaux J, Ìakar PZ, Bailey JE, Sauer U, Szyperski T: **Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional ^{13}C labelling of common amino acids.** *Eur J Biochem* 2001, **268:**2464-2479.
35. Nissen TL, Schulze U, Nielsen J, Villadsen J: **Flux distributions in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae*.** *Microbiology* 1997, **143:**203-218.
36. Sonderegger M, Jeppsson M, Hahn-Hägerdal B, Sauer U: **Molecular Basis for Anaerobic Growth of *Saccharomyces cerevisiae* on Xylose, Investigated by Global Gene Expression and Metabolic Flux Analysis.** *Appl Environ Microbiol* 2004, **70:**2307-2317.
37. Daran-Lapujade P, Rossell S, van Gulik WM, Luttik MAH, de Groot MJL, Slijper M, Heck AJR, Daran J-M, de Winde JH, Westerhoff HV, Pronk JT, Bakker BM: **The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels.** *PNAS* 2007, **104:**15753-15758.
38. van Winden W, van Dam JC, Ras C, Kleijn RJ, Vinke JL, van Gulik WM, Heijnen JJ: **Metabolic-flux analysis of *Saccharomyces cervisiae* CEN.PK113-7D based on mass isotopomer measurements of ^{13}C-labeled primary metabolites.** *FEMS Yeast Res* 2005, **5:**559-568.
39. Grosz R, Stephanopoulos G: **Physiological, Biochemical, and Mathematical Studies of Micro-Aerobic Continuous Ethanol Fermentation by *Saccharomyces cerevisiae* I: Hysteresis, Oscillations, and Maximum Specific Ethanol Productivities in Chemostat Culture.** *Biotechnol Bioeng* 1990, **36:**1006-1019.
40. Zitomer RS, Lowry CV: **Regulation of Gene Expression by Oxygen in *Saccharomyces cerevisiae*.** *Microbiol Rev* 1992, **56:**1-11.
41. Kwast KE, Burke PV, Boyton RO: **Oxygen sensing and the transcriptional regulation of oxygen-responsive genes in yeast.** *J Exp Biol* 1998, **201:**1177-1195.
42. van Dijken JP, Bauer J, Brambilla L, Dupoc P, Francois JM, Gancedo FC, Giuseppin MLF, Heijnen JJ, Hoare M, Lange HC, Madden EA, Niederberger P, Nielsen J, Parrou JL, Petit T, Porro D, Reuss M, van Riel N, Rizzi M, Steensma HY, Verrips CT, Vindeløv J, Pronk JT: **An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains.** *Enzyme Microb Technol* 2000, **26:**706-714.
43. Szyperski T: **Biosynthetically directed fractional ^{13}C-labelling of proteinogenic amino acids. An efficient tool to investigate intermediary metabolism.** *Eur J Biochem* 1995, **232:**433-448.
44. Kleijn RJ, van Winden WA, van Gulik WM, Heijnen JJ: **Revisiting the ^{13}C-label distribution of the non-oxidative branch of the pentose phosphate pathway based upon kinetic and genetic evidence.** *FEBS J* 2005, **272:**4970-4982.
45. Camarasa C, Grivet JP, Dequin S: **Investigation by ^{13}C-NMR and tricarboxylic acid (TCA) deletion mutant analysis of pathways for succinate formation in *Saccharomyces cerevisiae* during anaerobic fermentation.** *Microbiology* 2003, **149:**2669-2678.
46. Rosenfeld E, Beauvoit B, Rigoulet M, Salmon JM: **Non-respiratory oxygen consumption pathways in anaerobically-grown *Sac-***

***charomyces cerevisiae*: evidence and partial characterization.** *Yeast* 2002, **19:**1299-1321.
47. ***Saccharomyces* Genome Database** [http://www.yeastgenome.org/]
48. Flikweert MT, Zanden L van der, Janssen WMTM, Steensma HY, van Dijken JP, Pronk JT: **Pyruvate decarboxylase: An indispensable enzyme for growth of *Saccharomyces cerevisiae* on glucose.** *Yeast* 1996, **12:**247-257.
49. Semenza GL: **Oxygen-dependent regulation of mitochondrial respiration by hypoxia-inducible factor 1.** *Biochem J* 2007, **405:**1-9.
50. Fukuda R, Zhang H, Kim JW, Shimoda L, Dang CV, Semenza GL: **HIF-1 Regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells.** *Cell* 2007, **129:**111-122.
51. Pronk JT, Steensma HY, van Dijken JP: **Pyruvate metabolism in *Saccharomyces cerevisiae*.** *Yeast* 1996, **12:**1607-1633.
52. Bakker B, Bro C, Kötter P, Luttik MAH, van Dijken JP, Pronk JT: **The mitochondrial alcohol dehydrogenase Adh3p is involved in a redox shuttle in *Saccharomyces cerevisiae*.** *J Bacteriol* 2000, **182:**4730-4737.
53. Costenoble R, Valadi H, Gustafsson L, Niklasson C, Franzén CJ: **Microaerobic glycerol formation in *Saccharomyces cerevisiae*.** *Yeast* 2000, **16:**1483-1495.
54. Vanrolleghem PA, de Jong-Gubbels P, van Gulik WM, Pronk JT, van Dijken JP, Heijnen S: **Validation of a Metabolic Network for *Saccharomyces cerevisiae* Using Mixed Substrate Studies.** *Biotechnol Prog* 1996, **12:**434-448.
55. Rigoulet M, Leverve X, Fontaine E, Ouhabi R, Guérin B: **Quantitative analysis of some mechanisms affecting the yield of oxidative phosphorylation: Dependence upon both fluxes and forces.** *Mol Cell Biochem* 1998, **184:**35-52.
56. Avéret N, Fitton V, Bunoust O, Rigoulet M, Guérin B: **Yeast mitochondrial metabolism: From *in vitro* to *in situ* quantitative study.** *Mol Cell Biochem* 1998, **184:**67-79.
57. Ouhabi R, Rigoulet M, Guerin B: **Flux-yield dependence of oxidative phosphorylation at constant $\Delta\mu H^+$.** *FEBS Lett* 1989, **254:**199-202.
58. Vaseghi S, Baumeister A, Rizzi M, Reuss M: ***In vivo* Dynamics of the Pentose Phosphate Pathway in *Saccharomyces cerevisiae*.** *Metab Eng* 1999, **1:**128-140.
59. Bruinenberg PM, van Dijken JP, Scheffers WA: **A theoretical analysis of NADPH production and consumption in yeasts.** *J Gen Microbiol* 1983, **129:**953-964.
60. Minard KI, McAlister-Henn L: **Dependence of Peroxisomal β-Oxidation of Cytosolic Sources of NADPH.** *J Biol Chem* 1999, **274:**3402-3406.
61. Cakir T, Kirdar B, Önsan ZI, Ulgen KÖ, Nielsen J: **Effect of carbon source perturbations on transcriptional regulation of metabolic fluxes in *Saccharomyces cerevisiae*.** *BMC Syst Biol* 2007, **1:**18.
62. Daran-Lapujade P, Jansen ML, Daran JM, van Gulik W, de Winde JH, Pronk JT: **Role of transcriptional regulation in controlling fluxes in central carbon metabolism of *Saccharomyces cerevisiae*, a chemostat culture study.** *J Biol Chem* 2004, **279:**9125-9138.
63. De Groot M, Daran-Lapujade P, van Breukelen B, Knijnenburg T, de Hulster E, Reinders M, Pronk J, Heck A, Slijper M: **Quantitative proteomics and Transcriptomics of anaerobic and aerobic yeast cultures reveals post-transcriptional regulation of key cellular processes.** *Microbiology* 2007, **153:**3864-3878.
64. Elbing K, Ståhlberg A, Hohmann S, Gustafsson L: **Transcriptional responses to glucose at different glycolytic rates in *Saccharomyces cerevisiae*.** *Eur J Biochem* 2004, **271:**4855-4864.
65. de Jong-Gubbels P, Bauer J, Niederberger P, Stückrath I, Kötter P, van Dijken JP, Pronk JT: **Physiological characterisation of a pyruvate-carboxylase-negative *Saccharomyces cerevisiae* mutant in batch and chemostat.** *Antonie Van Leeuwenhoek* 1998, **74:**253-263.
66. Verduyn C, Postma E, Scheffers WA, van Dijken JP: **Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation.** *Yeast* 1992, **8:**501-517.
67. Sauer U, Hatzimanikatis V, Bailey JE, Hochuli M, Szyperski T, Wüthrich K: **Metabolic fluxes in riboflavin-producing *Bacillus subtilis*.** *Nat Biotechnol* 1997, **15:**448-452.
68. Sauer U, Lasko DR, Fiaux J, Hochuli M, Glaser R, Szyperski T, Wüthrich K, Bailey JE: **Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism.** *J Bacteriol* 1999, **181:**6679-6688.
69. Sola A, Maaheimo H, Ylonen K, Ferrer P, Szyperski T: **Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*.** *Eur J Biochem* 2004, **271:**2462-2470.
70. Lange HC: *Quantitative Physiology of S. cerevisiae using Metabolic Network Analysis* PhD thesis. Technical University Delft, Delft, The Netherlands; 2002.
71. Swiegers J, Dippenaar N, Pretorius I, Bauer F: **Carnitine-dependent metabolic activities in *Saccharomyces cerevisiae*: three carnitine acetyltranferases are essential in a carnitine-dependent strain.** *Yeast* 2001, **18:**585-595.
72. van Roermund C, Hettema E, Berg M van der, Tabak H, Wanders R: **Molecular characterization of carnitine-dependent transport of acetyl-CoA from peroxisomes to mitochondria in *Saccharomyces cerevisiae* and identification of a plasma membrane carnitine transporter, Agp2p.** *EMBO J* 1999, **18:**5843-5852.
73. Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, Schönfisch B, Perschil I, Chacinska A, Guiard B, Rehling P, Pfanner N, Meisinger C: **The proteome of *Saccharomyces cerevisiae* mitochondria.** *PNAS* 2003, **100:**13207-13212.
74. Palmieri L, Vozza A, Agrimi G, De Marco V, Runswick J, Palmieri F, Walker JE: **Identification of the Yeast Mitochondrial Transporter for Oxaloacetate and Sulfate.** *J Biol Chem* 1999, **274:**22184-22190.
75. Visser W, Baan AA van der, Batenburg-van der Vegte W, Scheffers A, Krämer R, van Dijken JP: **Involvement of mitochondria in the assimilatory metabolism of anaerobic *Saccharomyces cerevisiae* cultures.** *Microbiology* 1994, **140:**3039-3046.
76. Boles E, de Jong-Gubbels P, Pronk JT: **Identification and Characterization of *MAE1*, the *Saccharomyces cerevisiae* Structural Gene Encoding Mitochondrial Malic Enzyme.** *J Bacteriol* 1998, **180:**2875-2882.
77. Antoniewicz MR, Kelleher JK, Stephanopoulos G: **Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements.** *Metab Eng* 2006, **8:**324-337.

PUBLICATION II

Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates

In: Microbiology 2007. Vol. 153, pp. 281–290.

Microbiology (2007), **153**, 281–290 DOI 10.1099/mic.0.29263-0

Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates

Aina Solà,[1]† Paula Jouhten,[2]† Hannu Maaheimo,[2] Francesc Sánchez-Ferrando,[3] Thomas Szyperski[4] and Pau Ferrer[1]

Correspondence
Pau Ferrer
pau.ferrer@uab.cat

[1]Department of Chemical Engineering, Universitat Autònoma de Barcelona, 08193-Bellaterra, Spain

[2]NMR-laboratory, VTT Technical Research Centre of Finland, PO Box 65, FIN-00014 Helsinki, Finland

[3]Department of Chemistry, Universitat Autònoma de Barcelona, 08193-Bellaterra, Spain

[4]Department of Chemistry, University at Buffalo, The State University of New York at Buffalo, NY 14260, USA

Received 29 June 2006
Revised 19 September 2006
Accepted 9 October 2006

The metabolic pathways associated with the tricarboxylic acid cycle intermediates of *Pichia pastoris* were studied using biosynthetically directed fractional ^{13}C labelling. Cells were grown aerobically in a chemostat culture fed at two dilution rates (1.39×10^{-5} s^{-1} and 4.44×10^{-5} s^{-1}) with varying mixtures of glycerol and methanol as sole carbon sources. The results show that, with co-assimilation of methanol, the common amino acids are synthesized as in *P. pastoris* cells grown on glycerol only. During growth at the lower dilution rate, when both substrates are entirely consumed, the incorporation of methanol into the biomass increases as the methanol fraction in the feed is increased. Moreover, the co-assimilation of methanol impacts on how key intermediates of the pentose phosphate pathway (PPP) are synthesized. In contrast, such an impact on the PPP is not observed at the higher dilution rate, where methanol is only partially consumed. This finding possibly indicates that the distribution of methanol carbon into assimilatory and dissimilatory (direct oxidation to CO_2) pathways are different at the two dilution rates. Remarkably, distinct flux ratios were registered at each of the two growth rates, while the dependency of the flux ratios on the varying fraction of methanol in the medium was much less pronounced. This study brings new insights into the complex regulation of *P. pastoris* methanol metabolism in the presence of a second carbon source, revealing important implications for biotechnological applications.

INTRODUCTION

The methylotrophic yeast *Pichia pastoris* has emerged as an important production host for both industrial protein production and basic research, including structural genomics (Lin Cereghino & Cregg, 2000; Lin Cereghino *et al.*, 2002; Yokoyama, 2003; Prinz *et al.*, 2004). However, progress in strain improvement and rational design and optimization of culture conditions for heterologous protein production in *P. pastoris* is currently hampered by the limited number of systematic metabolic and physiological characterization studies under bioprocess-relevant conditions (Sauer *et al.*, 2004; Solà *et al.*, 2004). Information on heterologous gene expression and production of proteins under different physiological states of the cells is scarce. Furthermore, very little information is available on the cellular responses to protein production in *P. pastoris* (Hohenblum *et al.*, 2004). Importantly, the *P. pastoris* genome has been deciphered (see www.integratedgenomics.com), offering innumerable possibilities to pursue coordinated understanding of cellular processes in the framework of systems biology.

P. pastoris has been developed as an expression platform using elements that include strong inducible promoters derived from genes of the methanol utilization pathway, which is compartmentalized in the peroxisomes (Harder & Veenhuis, 1989). During growth on methanol, several key enzymes, e.g. alcohol oxidase, catalase, formaldehyde

†These authors contributed equally to this work.

Abbreviations: [^{13}C,^{1}H]-COSY, [^{13}C,^{1}H] correlation NMR spectroscopy; BDF, biosynthetically directed fractional; cyt, cytosolic; GCV, glycine cleavage pathway; mt, mitochondrial; OAA, oxaloacetate; PPP, pentose phosphate pathway; SHMT, serine hydroxymethyltransferase; TCA, tricarboxylic acid.

Tables of *f*-values are available as supplementary data with the online version of this paper.

dehydrogenase and dihydroxyacetone synthase, are present in high amounts and peroxisomes proliferate. The synthesis of these enzymes is regulated at the transcriptional level of the respective genes. Methanol assimilation is subject to a carbon-source-dependent repression/derepression/induction mechanism; it is rather strongly repressed by multicarbon sources such as glucose and glycerol, but highly induced by methanol. Importantly, co-assimilation of a multicarbon source and methanol can be triggered under certain growth conditions (Egli *et al.*, 1982).

Although pathways of methanol metabolism are essentially analogous for all methylotrophic yeasts, important variations do exist with respect to their regulation (Harder & Veenhuis, 1989). It is, for example, well documented that during aerobic growth of different species (e.g. *Hansenula polymorpha*), partial catabolite repression of methanol metabolism specific enzymes may occur, i.e. allowing expression of these enzymes to quite significant levels. Furthermore, control by catabolite repression by different multicarbon compounds in some strains is tighter than in others, while methanol may have a small or a significant inducing effect. In *P. pastoris*, high-level induction of methanol metabolism enzymes is strongly dependent on methanol, i.e. partial catabolite repression of methanol-metabolism-specific enzymes only occurs at a much reduced level. Implications of the regulation of methanol metabolism for central carbon metabolism in *P. pastoris* growing on mixtures of methanol plus a multicarbon source are essentially unexplored. Moreover, most of the comprehensive investigations of methanol mixed carbon metabolism have so far been pursued only for other methylotrophic yeasts (e.g. *H. polymorpha*, *Kloeckera*, *Candida boidinii*; for a review see Harder & Veenhuis, 1989). In this context, the level of protein expression in *P. pastoris* depends critically on the growth conditions, and the attainment of high cell densities has been shown to improve protein yields substantially (Stratton *et al.*, 1998). Although production of recombinant proteins under such culture conditions is typically induced by methanol, which activates the *aox-1* promoter controlling the heterologous gene, feeding mixtures of glycerol (or other multicarbon sources) to the culture has also been successfully used as a means for improving process productivities (for a review see Cos *et al.*, 2006). In view of the outstanding role of *P. pastoris* for biotechnology research, this organism represents an obvious target for studies of its metabolism and physiology.

Stable isotope labelling experiments employed in conjunction with NMR spectroscopy and/or mass spectrometry (Szyperski, 1998) are a powerful tool for metabolic studies. In particular, biosynthetically directed fractional (BDF) ^{13}C labelling of proteinogenic amino acids has been developed into a cost-effective approach to assess the topology of active bioreactions (i.e. active pathways) and to quantify metabolic flux ratios (Szyperski, 1995). BDF labelling has been applied to study central carbon metabolism of eubacteria (Szyperski, 1995; Sauer *et al.*, 1997, 1999) as well as eukaryotic yeast cells (Maaheimo *et al.*, 2001; Fiaux *et al.*, 2003; Solà *et al.*, 2004) growing on glucose or glycerol.

Recently, we have established BDF ^{13}C labelling and metabolic flux ratio formalism (Szyperski, 1995; Maaheimo *et al.*, 2001) as an analytical tool to study intermediary carbon metabolism of *P. pastoris* cells growing on glycerol as sole carbon source in chemostat cultures (Solà *et al.*, 2004). This investigation allowed accurate mapping of the metabolic state of the tricarboxylic acid (TCA) cycle and associated pathways, thus providing a valuable methodological basis for the analysis of *P. pastoris* cells growing on mixtures of glycerol and methanol, which is described in the present study. In addition, here we have applied the metabolic flux ratio formalism for yeast growing on a single carbon source (Maaheimo *et al.*, 2001) to the case of two-carbon-source co-assimilation.

METHODS

Strain and media. A prototrophic *P. pastoris* strain expressing a heterologous protein – a *Rhizopus oryzae* lipase (ROL) – under the transcriptional control of the *aox-1* promoter was chosen for metabolic flux ratio profiling. *P. pastoris* X-33/pPICZαA-ROL (Minning *et al.*, 2001) is the wild-type phenotype X-33 strain (Invitrogen) with the pPICZαA-derived expression vector (Invitrogen) containing the *ROL* gene, pPICZαA-ROL, integrated in its *aox-1* locus. Chemostat cultures were fed with a defined minimal medium containing (per 1×10^{-3} m^3 of deionized water): Yeast Nitrogen Base (YNB; Difco), 0.17×10^{-3} kg; $(NH_4)_2SO_4$, 5×10^{-3} kg; glycerol and methanol (different ratios on w/w basis), 10×10^{-3} kg (total); Antifoam Mazu DF7960 (Mazer Chemicals, PPG Industries), 0.1×10^{-6} m^3. The YNB components and methanol were sterilized separately by microfiltration and then added to the bioreactor. The medium used for starter cultures was YPD medium containing 1 % (w/v) yeast extract, 2 % (w/v) peptone, 2 % (w/v) glucose.

Chemostat cultures. Continuous cultures were carried out at a working volume of 0.8×10^{-3} m^3 in a 1.5×10^{-3} m^3 bench-top bioreactor (BiofloIII; New Brunswick) at 30 °C and with a minimum dissolved oxygen tension of 30 %. Simultaneous cultures using glycerol and methanol in different proportions as carbon source were performed at two different dilution rates, D (defined as volumetric flow rate/working volume) of 1.39×10^{-5} s^{-1} and 4.44×10^{-5} s^{-1}. These values are just below the maximum specific growth rate, μ_{max}, of *P. pastoris* cells growing on an excess of methanol or glycerol, 1.94×10^{-5} and 4.72×10^{-5} s^{-1}, respectively (Solà, 2004). Medium feeding was controlled by a Masterflex pump (Cole-Parmer). The working volume was kept constant by removal of effluent from the centre of the culture volume by use of a peristaltic pump (B. Braun Biotech). The pH of the culture was maintained at 5.5 by addition of 1 M KOH and the airflow was maintained at 0.16×10^{-4} m^3 s^{-1} with filter-sterilized air using a mass flow controller (Brooks Instruments). The agitation speed was set to 500 r.p.m. Starter cultures (1×10^{-4} m^3) were grown in 1 l baffled shake flasks at 200 r.p.m. at 30 °C for 8.64×10^4 s. Cells were harvested by centrifugation and resuspended in fresh medium prior to the inoculation of the bioreactor. The culture was initially run in batch mode to grow cells until the late exponential growth phase and then switched to continuous operational mode.

Analytical procedures. Cell biomass was monitored by measuring OD_{600}. For cellular dry weight, a known volume of culture broth was filtered using pre-weighed filters; these were washed with 2 vols

distilled water and dried to constant weight at 105 °C for 8.64×10^4 s. Samples for extracellular metabolite analyses were centrifuged at 6000 r.p.m. for 120 s in a microcentrifuge to remove the cells. Glycerol, acetic acid and ethanol were analysed by HPLC as described by Solà *et al.* (2004). Methanol was measured by GC as described by Minning *et al.* (2001). The exhaust gas of the bioreactor was cooled in a condenser at 2–4 °C (Frigomix R; B. Braun Biotech) and dried through a silica gel column. Concentrations of oxygen and CO_2 in the exhaust gas of bioreactor cultures were determined on line with a mass spectrometer (Omnistar; Balzers Instruments).

BDF ^{13}C labelling. *P. pastoris* cells were fed with a minimal medium containing 10 kg different glycerol/methanol mixtures m^{-3} (8 : 2, 6 : 4 and 4 : 6, w/w) for five volume changes to reach a metabolic steady-state, as indicated by a constant cell density and constant oxygen and CO_2 concentrations in the bioreactor exhaust gas. BDF ^{13}C labelling of cells growing at steady-state on a single carbon source has been described elsewhere (Sauer *et al.*, 1997; Fiaux *et al.*, 2003; Solà *et al.*, 2004); essentially, it is achieved by feeding the reactor with medium containing about 10 % (w/w) of uniformly ^{13}C-labelled and 90 % unlabelled substrate for one volume change. In this study, where two carbon sources (namely glycerol and methanol) were used simultaneously, the BDF ^{13}C labelling step involved feeding the reactor with medium containing about 10 % (w/w) uniformly ^{13}C-labelled and 90 % unlabelled amounts of each substrate simultaneously for one volume change. Uniformly ^{13}C-labelled glycerol (isotopic enrichment of >98 %) was purchased from Martek Biosciences or Spectra Stable Isotopes. ^{13}C-labelled methanol (isotopic enrichment of 99 %) from Cambridge Isotope Laboratories was purchased from Euriso-top. Cells were then harvested by centrifugation at 4000 ***g*** for 600 s, resuspended in 2×10^{-2} M Tris/HCl (pH 7.6) and centrifuged again. Finally, the washed cell pellets were lyophilized (Benchtop 5L Virtis Sentry), of which 2×10^{-4} kg were resuspended in 3×10^{-6} m^3 of 2×10^{-2} M Tris/HCl (pH 7.6). After addition of 6×10^{-6} m^3 6 M HCl, the biomass was hydrolysed in sealed glass tubes at 110 °C for 8.64×10^4 s, the solutions were filtered using 0.2 μm filters (Millex-GP; Millipore) and lyophilized.

NMR spectroscopy and data analysis. The lyophilized hydrolysates were dissolved in 0.1 M DCl in D_2O and two-dimensional (2D) [^{13}C,^{1}H] correlation NMR spectroscopy (COSY) spectra were acquired for both aliphatic and aromatic resonances as described previously (Szyperski, 1995) at 40 °C on a Varian Inova spectrometer operating at a ^{1}H resonance frequency of 600 MHz. The spectra were processed using standard Varian spectrometer software VNMR (version 6.1, C). The program FCAL (R. W. Glaser; FCAL 2.3.1) (Szyperski *et al.*, 1999) was used for the integration of ^{13}C–^{13}C scalar fine structures in 2D [^{13}C,^{1}H]-COSY, for the calculation of relative abundances, *f*-values, of intact carbon fragments arising from a single carbon source molecule (Szyperski, 1995), and for the calculation of the resulting flux ratios through several key pathways in central metabolism (Szyperski, 1995; Maaheimo *et al.*, 2001). The probabilistic equations relating the ^{13}C fine structures to *f*-values can be readily applied to this case of two simultaneous carbon sources. This is because, as a C_1-compound, methanol does not introduce contiguous multiple-carbon fragments to the metabolism and, therefore, all contiguous $^{13}C_n$ ($n>1$) fragments must originate from glycerol. Since the probabilistic equations for calculating the flux ratios depend on a uniform degree of ^{13}C labelling, both glycerol and methanol were supplied with the same fraction of uniformly ^{13}C-labelled molecules.

As described previously (Szyperski, 1995, 1998; Sauer *et al.*, 1997, 1999; Szyperski *et al.*, 1999; Maaheimo *et al.*, 2001; Fiaux *et al.*, 2003; Solà *et al.*, 2004), the calculation of metabolic flux ratios when using fractional ^{13}C labelling of amino acids is based on assuming both a metabolic (see above) and an isotopomeric steady-state. To establish a cost-effective protocol for a larger number of ^{13}C labelling experiments, we fed a chemostat operating in metabolic steady-state for the duration of one volume change with the medium containing the ^{13}C-labelled substrates (Sauer *et al.*, 1997; Fiaux *et al.*, 2003) before harvesting the biomass. Then, the fraction of unlabelled biomass produced prior to the start of the supply with ^{13}C-labelled medium can be calculated following simple wash-out kinetics (Szyperski, 1998; see also Solà *et al.*, 2004 for additional discussion).

Measurement of the degree of ^{13}C enrichment in CO_2. For the determination of ^{13}C incorporation from ^{13}C-labelled methanol to CO_2, cells were first cultivated with unlabelled medium containing a given glycerol/methanol mixture as carbon source until steady-state was achieved, as described above. During one residence time at steady-state, the CO_2 produced was trapped by bubbling the outlet air through a tube containing 2×10^{-5} m^3 of 10 M KOH. The culture was then fed with medium containing about 50 % (w/w) uniformly ^{13}C-labelled and 50 % unlabelled methanol plus unlabelled glycerol at the same ratio as in the unlabelled medium for one volume change. The $^{13}CO_2$ produced was trapped by bubbling the outlet air through a tube containing 2×10^{-5} m^3 of 10 M KOH for the period of one residence.

The ^{13}C content of carbonate anions in culture off-gas samples was measured by ^{13}C NMR spectroscopy on a Bruker 500 Avance spectrometer using a cryoprobe to improve the signal to noise ratio. Samples were prepared by mixing 0.2×10^{-6} m^3 of the corresponding 10 M KOH solution with 0.2×10^{-6} m^3 of a 1 M solution of dioxane (internal standard for both calibration and integration) in D_2O. ^{13}C NMR spectra were obtained at 125 MHz for each sample under Waltz-16 proton decoupling, using a 31 450 Hz (over 250 p.p.m.) sweep width, with a 30 degree ^{13}C pulse and a relaxation delay of 1.0 s. After accumulation to a good signal to noise ratio, the flame ionization detectors were weighted with a 1.0 Hz line broadening function and Fourier transformed. The resulting spectra showed peaks at 166.6 p.p.m. (carbonate anion) and 66.9 p.p.m. (dioxane), which were integrated. ^{13}C incorporated into CO_2 was estimated by comparing the ^{13}C content in carbonate anions in corresponding unlabelled and labelled samples.

Biochemical reaction network model for *P. pastoris*. The biochemical reaction network model for data interpretation was the one recently identified for *Saccharomyces cerevisiae* (Maaheimo *et al.*, 2001; Fiaux *et al.*, 2003), which was also shown to be suitable for *Pichia stipitis* (Fiaux *et al.*, 2003) and *P. pastoris* (Solà *et al.*, 2004). Considering published data (Harder & Veenhuis, 1989), pathways for methanol metabolism were added (Fig. 1). Briefly, methanol is oxidized by an alcohol oxidase to generate formaldehyde, which is further oxidized to CO_2 or assimilated into carbohydrates. The first step in the formaldehyde assimilation pathway involves a dihydroxyacetone synthase, which catalyses the condensation of formaldehyde with xylulose 5-phosphate to form fructose 6-phosphate. The hydrogen peroxide formed in the initial oxidation of methanol is removed by the action of a catalase. These four enzymes are peroxisomal. Furthermore, methanol assimilation by yeasts is characteristically associated with the biogenesis of peroxisomes.

RESULTS AND DISCUSSION

P. pastoris cultures were performed at two dilution rates, 1.39×10^{-5} s^{-1} and 4.44×10^{-5} s^{-1}, in aerobic chemostats using mixtures of glycerol and methanol at different ratios as sole carbon sources. The lower dilution rate is slightly below the μ_{max} of the organism as observed previously in a batch culture on methanol (1.94×10^{-5} s^{-1}), i.e. where the

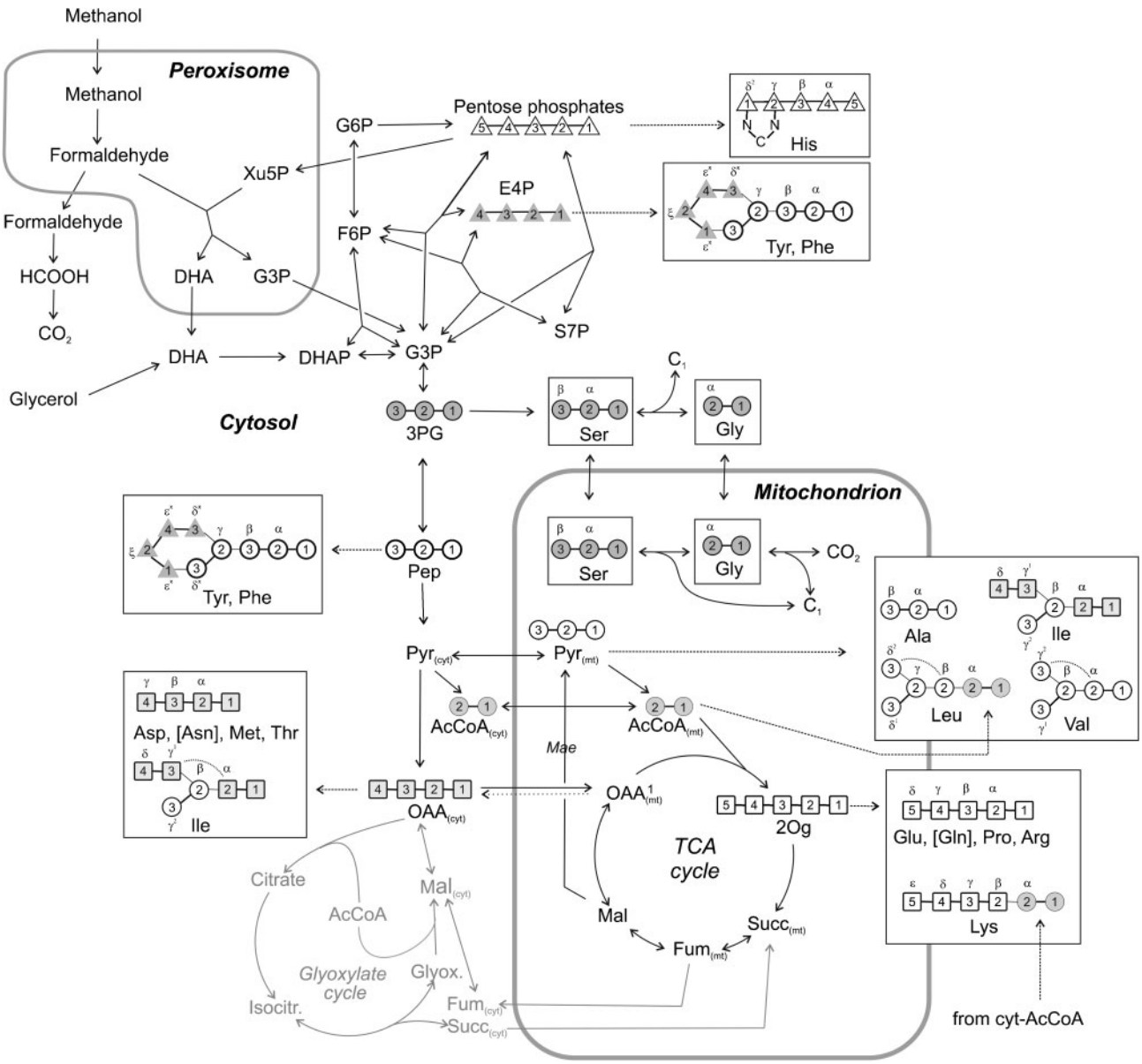

Fig. 1. Network of active biochemical pathways constructed for *P. pastoris* cells grown with glycerol and methanol as mixed carbon source. The network is based on those identified for *P. pastoris* growing on glucose or glycerol (Solà *et al.*, 2004) and on the literature on methanol metabolism of methylotrophic yeasts (Harder & Veenhuis, 1989; see text). The central carbon metabolism of *P. pastoris* is dissected into cytosolic and mitochondrial subnetworks. In addition, the reactions involved in the initial oxidation steps of methanol to formaldehyde (i.e. alcohol oxidase and catalase), the first reaction involved in formaldehyde fixation (i.e dihydroxyacetone synthase), as well as the glyoxylate cycle reactions are supposed to reside in peroxisomes in methylotrophic yeast like *P. pastoris*. Since the reactions of the glyoxylatecycle cannot be identified with the current ^{13}C labelling strategy (see text), its reactions are depicted in grey. Amino acids and carbon fragments originating from a single intermediate of central carbon metabolism are represented in the rectangular boxes. Thin lines between amino acid carbon atoms denote carbon bonds that are formed between fragments originating from different precursor molecules, while thick lines indicate intact carbon connectivities in fragments arising from a single precursor molecule. The carbon skeletons of glycolysis, TCA cycle and PPP intermediates are represented by circles, squares and triangles, respectively. The numbering of the carbon atoms refers to the corresponding atoms in the precursor molecule. Abbreviations: AcCoA, acetyl-Coenzyme A; DHA, dihydroxyacetone; DHAP, dihydroxyacetone phosphate; E4P, erythrose 4-phosphate; F6P, fructose 6-phosphate; Fum, fumarate; G6P, glucose 6-phosphate; Gly, glycine; Glyox., glyoxylate; G3P, glyceraldehyde 3-phosphate; 3PG, 3-phosphoglycerate; Mae, malic enzyme; Mal, malate; OAA, oxaloacetate; 2Og, 2-oxoglutarate; Pyr, pyruvate; Pep, phosphoenolpyruvate; S7P, sedoheptulose-7-phosphate; Ser, serine; Succ, succinate; Xu5P, xylulose 5-phosphate. For AcCoA, Fum, OAA, Pyr and Succ, cytosolic (cyt) and mitochondrial (mt) pools are indicated separately.

glycerol supply is growth-limiting. The higher dilution rate is slightly below the μ_{max} of the organism observed previously in a batch culture on glycerol (4.72×10^{-5} s^{-1}) (Table 1).

All chemostat cultures operated at $D = 1.39 \times 10^{-5}$ s^{-1} simultaneously utilized glycerol and methanol, indicating that glycerol repression of methanol consumption did not occur. In fact, the residual concentrations of glycerol and methanol in the culture media were below the detection limits of the respective assays.

At the higher rate of $D = 4.44 \times 10^{-5}$ s^{-1} with an 80 : 20 glycerol/methanol mixture, some residual glycerol (2.2 kg glycerol m^{-3}) accumulated in the growth medium and, concomitantly, very little methanol was consumed under these conditions. However, residual glycerol concentrations in the chemostat were very close to or below the detection limit when the glycerol/methanol ratio was decreased. Under such conditions, a significant fraction of the methanol was consumed by the cells, though the residual methanol concentration increased as the fraction of methanol increased. These results confirm that cells fed with mixtures of methanol and glycerol are able to utilize methanol at dilution rates considerably higher than μ_{max} in batch cultures grown on methanol as sole carbon source ($D = 1.94 \times 10^{-5}$ s^{-1}) (Zhang *et al.*, 2003). A similar substrate utilization pattern has been observed in *H. polymorpha* growing on different methanol/glucose mixtures (ranging from 0 : 100 to 100 : 0) and growth rates (Egli *et al.*, 1986). At low dilution rates both carbon sources were utilized simultaneously, but at higher dilution rates the cells increasingly accumulated methanol in the culture medium. The dilution rate at which the transition from glucose/methanol growth to glucose growth occurred (D_t) was strictly dependent on the composition of the methanol/glucose mixture in the feed, and D_t increased with decreasing proportions of methanol. Similarly, growth of *P. pastoris* at $D = 4.44 \times 10^{-5}$ s^{-1} is probably close to the upper limit of the specific growth rate at which the regulatory mechanism that determines the onset of repression of methanol-assimilating enzymes in cells growing on glycerol mixtures.

Notably, ethanol and acetate were not detected by HPLC in any of the cultures, and carbon balances closed within 5 %. Hence, *P. pastoris* cells, when growing under the experimental conditions described, used both glycerol and methanol entirely to generate biomass and CO_2. The observed biomass yields ($Y_{x/s}$) in these mixed-substrate cultures gave a reasonable fit with the predicted $Y_{x/s}$ calculated as the weighted mean of the growth yields on the two individual substrates (Table 1). These were calculated from an aerobic chemostat culture at $D = 1.39 \times 10^{-5}$ s^{-1} using methanol as sole carbon source [0.31 kg cell dry wt (kg glycerol)$^{-1}$; Solà, 2004] and from chemostat cultures at $D = 1.39 \times 10^{-5}$ s^{-1} and 4.44×10^{-5} s^{-1} using glycerol as sole carbon source [0.63 kg cell dry wt (kg methanol)$^{-1}$; Solà *et al.*, 2004]. An analogous pattern has been observed in chemostat cultures of *H. polymorpha* growing on different glucose/methanol mixtures (Egli *et al.*, 1986). Also, during growth at $D = 1.39 \times 10^{-5}$ s^{-1} and 4.44×10^{-5} s^{-1} both the specific methanol consumption rate (q_{met}) and specific CO_2 production rate (q_{CO_2}) increased proportionally as the glycerol/methanol ratio decreased. However, this does not necessarily imply that no change in distribution of methanol carbon into assimilatory and dissimilatory pathways took place because of the presence of the second growth substrate, glycerol. Metabolic flux ratio analyses were performed with hydrolysed biomass samples that were harvested from these chemostat cultures in physiological steady-state. 2D [^{13}C,^{1}H]-COSY data were analysed as described by Maaheimo *et al.* (2001), yielding the desired relative abundances (*f*-values) of intact carbon

Table 1. Growth parameters in steady-state chemostat cultures of *P. pastoris*

$Y_{x/s}$ represents the biomass yield, q_{glyc}, q_{gluc} and q_{O_2} are specific utilization rates, q_{CO_2} is the specific production rate, Glyc and Meth indicate glycerol and methanol, respectively, and RQ is the respiratory quotient. ND, Not determined.

Carbon source	Residual substrate concn (Glyc/Meth; kg m^{-3})	$Y_{x/s}$ (kg dry wt kg^{-1}]	q_{glyc}/q_{meth}	q_{CO_2} (mol kg^{-1} per 3600 s)	q_{O_2} (mol kg^{-1} per 3600 s)	RQ
$D = 1.39 \times 10^{-5}$ s^{-1}						
Glycerol	0.0/–	0.63	1.09/–	1.56	2.16	0.72
80 glycerol/20 methanol	0.0/0.09	0.51	0.95/0.63	1.70	2.70	0.63
60 glycerol/40 methanol	0.0/0.17	0.44	0.74/1.48	2.10	3.90	0.54
40 glycerol/60 methanol	0.0/0.09	0.44	0.57/2.33	2.21	4.85	0.46
$D = 4.44 \times 10^{-5}$ s^{-1}						
Glycerol	3.0/–	0.63	2.75/–	2.35	3.62	0.65
80 glycerol/20 methanol	2.2/1.8	0.65	ND	ND	ND	ND
60 glycerol/40 methanol	0.05/2.6	0.51	2.77/1.87	4.18	7.19	0.58
40 glycerol/60 methanol	0.0/3.9	0.53	2.23/2.73	3.60	7.20	0.50

fragments arising from a single source molecule of glycerol (Tables S1 and S2, available with the online version of this paper).

Biosynthesis of proteinogenic amino acids and C1 metabolism in *P. pastoris*

As expected, the *f*-values obtained for the mixed glycerol/methanol cultures (Tables S1 and S2, available with the online version of this paper) show that proteinogenic amino acids are primarily synthesized in *P. pastoris* according to the pathways documented for *S. cerevisiae* (Jones & Fink, 1982; Voet & Voet, 1995; Michal, 1998; Maaheimo *et al.*, 2001), and these have also been validated for *P. pastoris* cells growing on glucose and glycerol (Solà *et al.*, 2004). Remarkably, the fraction of methanol in the feed affects the pool of Ser molecules effected by the reversible cleavage by serine hydroxymethyltransferase (SHMT); about 40–43 % of Ser molecules are cleaved in glucose and glycerol cultures (Solà *et al.*, 2004), but this fraction changes in mixed glycerol/methanol cultures. For instance, it decreased to about 28 % in cells growing at $D = 1.39 \times 10^{-5}$ s^{-1} in the 80 : 20 and 40 : 60 glycerol/methanol mixtures (Table 2). For Gly synthesis, yeasts can cleave either Ser (via SHMT) or Thr (via threonine aldolase). Due to the near degeneracy of *f*-values, however, it is not possible to accurately determine the relative contribution of the two pathways, or to distinguish between cytosolic and mitochondrial SHMT activity (Solà *et al.*, 2004). In contrast to the SHMT pathway, the Thr cleavage reaction via threonine aldolase is, if present,

Table 2. Origins of metabolic intermediates during aerobic growth of *P. pastoris* in glycerol/methanol chemostat cultures

For comparison, corresponding data reported previously for *P. pastoris* growing on glycerol in chemostat aerobic cultures (Solà *et al.*, 2004) are given in the left-most column. Glyc and Meth indicate glycerol and methanol, respectively. PEP, Phosphoenolpyruvate; PYR, pyruvate.

Metabolite	Fraction of total pool (mean % ± SD)			
	Glyc*	80 Glyc/20 Meth	60 Glyc/40 Meth	40 Glyc/60 Meth
$D = 1.39 \times 10^{-5}$ s^{-1}				
Cytosol				
PEP from cyt-OAA (PEP carboxykinase reaction)	<3	<7	9±6	<11
cyt-OAA from cyt-PYR†	32±2	37±4	35±2	33±3
cyt-OAA reversibly converted to fumarate at least once (cytosolic or inter-compartmental exchange)	56±13	50±16	58±19	48±25
Mitochondria				
mt-PYR from malate (upper bound)	<4	<9	<16	<19
mt-PYR from malate (lower bound)	<3	<2	<4	<3
mt-OAA from PEP (anaplerotic supply of TCA cycle)	33±2	32±3	36±3	29±4
mt-OAA reversibly converted to fumarate at least once	65±14	52±17	59±15	50±20
C1 metabolism				
Ser from Gly and C1 unit	43±3	28±2	37±3	28±2
Gly from CO_2 and C1 unit	2±2	4±3	9±3	5±2
cyt-Gly from mt-Gly	<2	<4	<9	<5
$D = 4.44 \times 10^{-5}$ s^{-1}				
Cytosol				
PEP from cyt-OAA (PEP carboxykinase reaction)	<6	<21	<10	<11
cyt-OAA from cyt-PYR†	68±4	89±2	78±2	76±2
cyt-OAA reversibly converted to fumarate at least once (cytosolic or inter-compartmental exchange)	12±6	<5	11±5	<6
Mitochondria				
mt-PYR from malate (upper bound)	<11	<7	<6	0
mt-PYR from malate (lower bound)	<3	<2	<2	0
mt-OAA from PEP (anaplerotic supply of TCA cycle)	48±2	47±2	42±2	43±2
mt-OAA reversibly converted to fumarate at least once	61±14	55±8	52±12	53±12
C1 metabolism				
Ser from Gly and C1 unit	42±2	37±2	33±2	51±2
Gly from CO_2 and C1 unit	5±3	3±3	4±2	7±2
cyt-Gly from mt-Gly	<5	<3	<4	<7

*Data taken from Solà *et al.* (2004).

†Values assuming absence of cytosolic OAA from fumarate conversion.

irreversible in all cultures. This can be readily deduced from the fact that nearly identical *f*-values were obtained from Thr and Asp. Gly may also be synthesized from a C1 unit and CO_2 via the mitochondrial glycine cleavage (GCV) pathway. In contrast to previous studies with *S. cerevisiae* (Maaheimo *et al.*, 2001) and with *P. pastoris* growing on glucose (Solà *et al.*, 2004), we found no evidence for efflux of Gly which had been reversibly cleaved by the GCV into the cytosol in glycerol and glycerol/methanol cultures. Hence, it may be that either the mitochondrial GCV pathway is operating irreversibly, or Gly is not exported into the cytosol when cells are grown on glycerol. Only when cells are grown at $D=1.39\times10^{-5}$ s^{-1} with a 60 : 40 glycerol/methanol mixture do the differences observed in the labelling patterns of Gly Cα and Phe Cα indicate that the mitochondrial GVC is operative. In principle, yeasts can also synthesize Gly from TCA cycle intermediates via isocitrate lyase and the alanine/glyoxylate aminotransferase (Takada & Noguchi, 1985). However, our data suggest that the activity of the glyoxylate cycle is low (see below), so that this route for Gly synthesis is probably of minor importance, if it is active at all.

Comparative flux ratio profiling of *P. pastoris* growing on glycerol- and methanol-limited mixtures in chemostats

The use of the three-carbon source glycerol and methanol for BDF ^{13}C labelling of proteinogenic amino acids enabled the determination of the flux ratios for reactions associated with the TCA cycle (Table 2), while those related to glycolysis and the pentose phosphate pathway (PPP) could not be assessed (Solà *et al.*, 2004). This is because when labelled glycerol is metabolized through gluconeogenesis and oxidative PPP, labelling patterns that are sufficiently distinct from those generated when glycerol is channelled through the non-oxidative PPP are not produced. In fact, the only information that can be derived with respect to the operation of the PPP is obtained from the *f*-values of His Cβ (Tables S1 and S2, available with the online version of this paper). The *f*-values reveal the reversible activity of the transketolase and transaldolase reactions when *P. pastoris* is grown on glycerol. Also, important variations in the observed E4P and R5P labelling patterns can be detected as the methanol fraction in the feed is increased (Table 3), suggesting an increasing activity of the methanol assimilation pathway, which involves PPP intermediates (Fig. 1). The increasing contribution of methanol to biomass constituents is further confirmed by the observation that the fraction of intact PEP molecules (i.e. originating from a single glycerol molecule) sharply decreases as the methanol fraction in the feed is increased, whereas the fraction of PEP molecules with both ^{13}C-^{13}C bonds cleaved (i.e. exclusively originating from methanol) increases (Table 3). Nevertheless, the fraction of methanol carbons assimilated by the cells that enter central carbon metabolism in relation to the methanol carbons that are dissimilated directly into CO_2 cannot be determined.

Ample information for pathways associated with TCA cycle intermediates can be obtained (Fig. 2) and can be summarized as follows. During growth at $D=1.39\times10^{-5}$ s^{-1} one finds that (i) gluconeogenesis from cytosolic oxaloacetate (cyt-OAA) via phosphoenolpyruvate (PEP) carboxykinase is either not detected or active at very low levels; (ii) synthesis of mitochondrial pyruvate (mt-PYR) from malate via malic enzyme is not detected; (iii) the

Table 3. Labelling patterns of erythrose 4-phosphate (E4P) and phosphoenolpyruvate (PEP) metabolic intermediates during aerobic growth of *P. pastoris* in glycerol/methanol chemostat cultures

For comparison, corresponding data reported previously for *P. pastoris* growing on glycerol in chemostat aerobic cultures (Solà *et al.*, 2004) are given in the left-most column. Glyc and Meth indicate glycerol and methanol, respectively. RSP, ribose 5-phosphate.

Metabolite	Fraction of total pool (%)			
	Glyc*	80 Glyc/20 Meth	60 Glyc/40 Meth	40 Glyc/60 Meth
$D=1.39\times10^{-5}$ s^{-1}				
Cytosol				
E4P with intact C_2-C_3-C_4 bonds	44	30	28	19
R5P with intact C_3-C_4-C_5 bonds	96	72	61	46
R5P with cleaved C_1-C_2 bond (TA/TK activity)	55	70	78	90
PEP with intact C_1-C_2-C_3 bonds	96−93	70−67	63	49−48
$D=4.44\times10^{-5}$ s^{-1}				
Cytosol				
E4P with intact C_2-C_3-C_4 bonds	46	43	45	44
R5P with intact C_3-C_4-C_5 bonds	94	86	95	89
R5P with cleaved C_1-C_2 bond (TA/TK activity)	55	65	58	55
PEP with intact C_1-C_2-C_3 bonds	97−93	90−89	97	91

*Data from Solà *et al.* (2004).

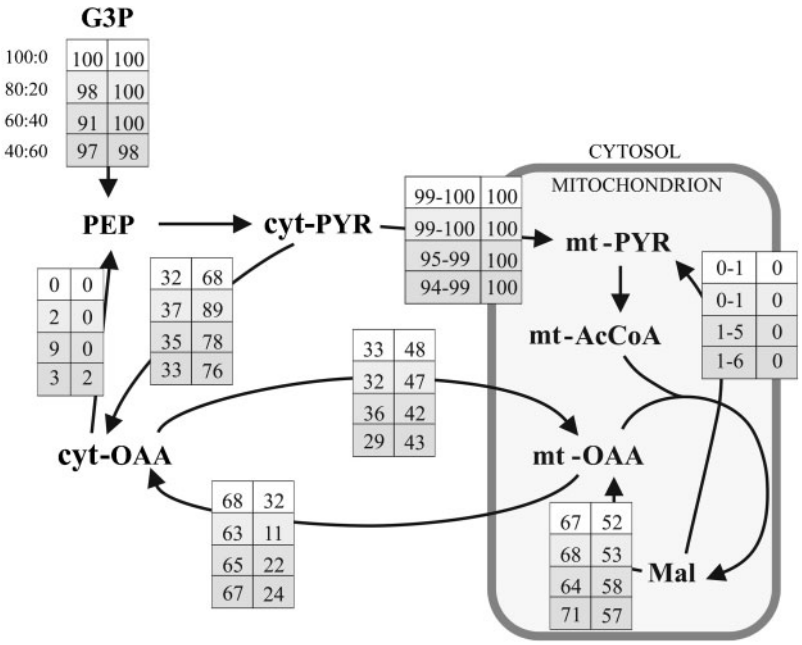

Fig. 2. Summary of flux information involving pools of TCA intermediates when *P. pastoris* cells are grown in a chemostat at $D = 1.39 \times 10^{-5}$ s^{-1} (left-hand values) and at $D = 4.44 \times 10^{-5}$ s^{-1} (right-hand values). The values in the boxes correspond from top to bottom, respectively, to the reference glycerol culture (data from Solà *et al.*, 2004), the 80 : 20, 60 : 40 and 40 : 60 mixed glycerol/methanol cultures. Note that values associated with arrows pointing at the same metabolite pool add up to 100 %. For abbreviations, see the legend to Fig. 1.

fraction of cyt-OAA that stems from the mitochondrial pool of C4 intermediates, e.g. via malate-Asp and/or malate-OAA shuttles (Bakker *et al.*, 2001), or possibly synthesized via the glyoxylate cycle, does not vary significantly as the fraction of methanol in the feed is increased; (iv) the fraction of mt-OAA reversibly interconverted to fumarate and the fraction of cyt-OAA reversibly interconverted to fumarate does not change significantly when the glycerol/methanol ratio is increased; (v) no significant alterations can be identified for the anaplerotic supply of the TCA cycle in cells growing with different glycerol/methanol ratios. The labelling patterns of cyt-OAA and mt-OAA molecules were almost identical, indicating relatively fast exchange between the two pools. Therefore, the anaplerosis was defined here as the fraction of mt-OAA molecules originating from PEP, which can be regarded as the flux of three-carbon molecules from glycolysis to the combined pool of cyt-OAA and mt-OAA.

Comparative flux ratio profiling of *P. pastoris* growing at high growth rates on glycerol/methanol mixtures in chemostats

At $D = 4.44 \times 10^{-5}$ s^{-1}, *P. pastoris* cells grow at about 90 % of the μ_{max} for glycerol provided as sole carbon source. Under these conditions, only part of the methanol being fed is effectively consumed by the cells. In fact, the residual methanol concentration in the culture broth increases with the methanol fraction in the feed, while the residual glycerol concentration is very low or zero (Table 1). In these cultures, the absolute amount of methanol effectively assimilated by the cells increases with the methanol fraction in the feed. Only in the case when glycerol accumulates to significant levels (80 : 20 glycerol/methanol mixture) is methanol virtually not consumed (Table 1). The comparison of flux ratios at this dilution rate revealed that, regardless of the methanol fraction in the feed medium, flux ratios are very similar to those observed in cells growing solely on glycerol as carbon source, e.g. the fraction of intact PEP molecules and the labelling patterns in PPP intermediates is not altered significantly by the assimilated methanol (Tables 2 and 3). *A priori*, this observation might indicate that the methanol that is consumed by the cells has a completely different fate from that in cells growing at the lower growth rate, namely that methanol could be mostly dissimilated directly to CO_2, generating 2 NADH molecules per methanol molecule. However, flux ratios in relation to the anaplerotic supply of the TCA cycle (i.e. the relative TCA cycle activity) are not significantly altered, as one would expect as a result of the extra amount of NADH produced by direct methanol oxidation to CO_2. Also, the observed ratio of the biomass to assimilated substrate ($Y_{x/s}$) still gave a reasonable fit with the predicted $Y_{x/s}$ calculated as the weighted mean of the growth yields on the two individual substrates (i.e. no significant drop in the observed growth yield was observed). Thus, the lower impact of methanol assimilation on the flux ratio distributions in cells growing at $D = 4.44 \times 10^{-5}$ s^{-1} may just reflect the fact that the amount of methanol molecules actually being consumed by the cells is significantly smaller than in cells growing at $D = 1.39 \times 10^{-5}$ s^{-1}, rather than a change in the split flux ratio between methanol assimilatory and dissimilatory pathways. In fact, a replica labelling experiment in which ^{13}C-labelled methanol (isotopic enrichment of 50 %) and unlabelled glycerol were fed at the 60 : 40 glycerol/methanol ratio indicated that the degree of enrichment of CO_2 was about 12.4 % at $D = 1.39 \times 10^{-5}$ s^{-1}, whereas incorporation

of ^{13}C into CO_2 at $D=4.44\times10^{-5}$ s^{-1} was close to negligible, i.e. supporting the latter hypothesis.

Comparison of flux ratio distributions of *P. pastoris* cells growing at $D=1.39\times10^{-5}$ s^{-1} on different glycerol/methanol mixtures with those observed in cells growing at $D=4.44\times10^{-5}$ s^{-1} on the same corresponding substrate mixtures shows the same pattern observed previously when comparing *P. pastoris* cells growing on glycerol as sole carbon source at these growth rates (Solà *et al.*, 2004). In particular, as also observed in cells growing on glycerol as sole carbon source, (i) the values for the fraction of cyt-OAA that stems from the mitochondrial pool of C4 intermediates, e.g. via malate-Asp and/or malate-OAA shuttles (Bakker *et al.*, 2001), or possibly synthesized via the glyoxylate cycle, are about twice as high at the lower dilution rate compared to the corresponding values in cells growing at the higher dilution rate. This indicates that, regardless of the fraction of methanol present in the feed medium, a largely unidirectional flux of OAA from the cytosol to the mitochondria occurs when cells are growing close to their maximal growth rates. (ii) The relative TCA cycle activity is reduced at the higher dilution rate (i.e. the fraction of mt-OAA generated by anaplerosis is a function of the specific growth rate but not of the methanol fraction in the feed medium) – unlike at the lower dilution rate, cyt-OAA and mt-OAA had distinct labelling patterns at the higher dilution rate, indicating slower exchange between the two pools. (iii) The fraction of cyt-OAA reversibly converted to fumarate is very low or zero in cells growing at the higher dilution rate, whereas this value is around 50 % in cells growing at the lower dilution rate. Blank & Sauer (2004) have defined this flux ratio as an upper limit of the fraction of cyt-OAA generated via the glyoxylate cycle, since export of TCA cycle intermediates from the mitochondria can also contribute to the cyt-OAA pool. However, the labelling patterns calculated for mt-OAA (as described by Maaheimo *et al.*, 2001) did not reflect a clear contribution of cytosolic succinate generated by the glyoxylate cycle.

Conclusions

This is the first comprehensive study of central carbon metabolism of the yeast *P. pastoris* growing on glycerol and methanol mixtures. In the framework of this study, we have established the BDF ^{13}C labelling approach of proteinogenic amino acids as an analytical tool to study intermediary metabolism of yeast cells growing on such carbon substrate mixtures. This approach allows the mapping of the metabolic state of the TCA cycle and associated pathways and thus this is an important methodological expansion for investigating the metabolism of eukaryotic cells growing with sole carbon sources. Specifically, we have shown that (i) co-assimilation of methanol as a carbon source does not alter the way the common amino acids are synthesized in *P. pastoris* growing on a sole multicarbon source, and (ii) growth on different glycerol/methanol mixtures at a given growth rate results in rather similar flux ratio profiles in the TCA cycle and related pathways as the fraction of methanol is increased. In contrast, a clear effect of specific growth rate on the relative activity of the TCA cycle and related pathways is observed, regardless of the methanol fraction in the feed, consistent with the observation that TCA cycle activity in *S. cerevisiae* is strongly correlated with the environmentally determined specific growth rate (Blank & Sauer, 2004).

Co-assimilation of methanol as a carbon source has a clear impact with respect to the activity of the PPP, which is consistent with the increasing flux of methanol molecules towards the synthesis of central carbon metabolism intermediates (e.g. PEP), as observed when the methanol fraction in the feed medium is increased. However, this pattern is not observed in cells growing at the higher dilution rate (where methanol is partially accumulated in the medium) suggesting that the distribution of methanol carbon into assimilatory and dissimilatory pathways may be different. Earlier ^{13}C labelling studies of methanol metabolism of the methylotrophic yeast *H. polymorpha* (Jones & Bellion, 1991) showed that the linear methanol oxidation pathway to CO_2 only operates under extreme conditions (e.g. methanol accumulation to toxic levels), suggesting a role in detoxification. Although the data obtained in the present study do not allow directly quantification of the split ratio of formaldehyde between the assimilation pathway and the oxidation pathway (Fig. 1) over the different tested environmental conditions, net fluxes through the metabolic network may be deduced from metabolic flux ratio analysis when combined with metabolic flux balancing (Fischer *et al.*, 2004; Fredlund *et al.*, 2004). Hence, we expect this study will lead to important insights into central carbon metabolism and its regulation in *P. pastoris*.

Overall, our investigation can be expected to become a valuable knowledge base for the optimization of culture processes for the production of recombinant proteins in *P. pastoris*, where parameters such as the residual methanol concentration, specific growth rate, as well as mixed substrate culture strategies have been shown to have a dramatic impact on overall process productivity. In addition, the information derived from our studies may be relevant for the design of isotopic labelling experiments of recombinant proteins (or other cell components, e.g. cell wall glucans) for structural studies. Furthermore, the methodology used in this work can also be applied to study the effect of other bioprocess-relevant parameters such as temperature, oxygen availability, etc., on the metabolic activity of *P. pastoris*.

ACKNOWLEDGEMENTS

This work was supported by the State University of New York at Buffalo, the Spanish Ministry of Science and Education (projects PPQ2001-1908 and CTQ2004-00300) and the Academy of Finland (projects 52311, 202409 and 207435). The authors thank O. Cos for technical assistance with culture off-gas analyses, as well as S. Santos and J. Albiol for physiological data consistency analysis. The

Department of Chemical Engineering of the Universitat Autònoma de Barcelona constitutes the Biochemical Engineering Unit of the Reference Centre in Biotechnology of the Generalitat de Catalunya.

REFERENCES

Bakker, B. M., Overkamp, K. M., van Maris, A. J. A., Kötter, P., Luttik, M. A. H., van Dijken, J. P. & Pronk, J. T. (2001). Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* **25**, 15–37.

Blank, L. M. & Sauer, U. (2004). TCA cycle activity in *Saccharomyces cerevisiae* is a function of the environmentally determined specific growth and glucose uptake rates. *Microbiology* **150**, 1085–1093.

Cos, O., Ramon, R., Montesinos, J. L. & Valero, F. (2006). Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: a review. *Microb Cell Fact* **5**, 17.

Egli, T., Käppeli, O. & Fiechter, A. (1982). Mixed substrate growth of methylotrophic yeasts in chemostat culture: influence of the dilution rate on the utilization of a mixture of glucose and methanol. *Arch Microbiol* **131**, 8–13.

Egli, T., Bosshard, C. & Hammer, G. (1986). Simultaneous utilization of methanol-glucose mixtures by *Hansenula polymorpha* in chemostat: influence of dilution rate and mixture composition on utilization pattern. *Biotechnol Bioeng* **28**, 1735–1741.

Fiaux, J., Çakar, Z. P., Sonderegger, M., Wüthrich, K., Szyperski, T. & Sauer, U. (2003). Metabolic flux profiling of the yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*. *Eukaryot Cell* **2**, 170–180.

Fischer, E., Zamboni, N. & Sauer, U. (2004). High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived ^{13}C constraints. *Anal Chem* **325**, 308–316.

Fredlund, E., Blank, L. M., Schnürer, J., Sauer, U. & Passoth, V. (2004). Oxygen- and glucose-dependent regulation of central carbon metabolism in *Pichia anomala*. *Appl Environ Microbiol* **70**, 5905–5911.

Harder, W. & Veenhuis, M. (1989). Metabolism of one-carbon compounds. In *The Yeasts, Vol. 3, Metabolism and Physiology of Yeasts*, pp. 289–316. Edited by A. H. Rose & J. S. Harrison. London: Academic Press.

Hohenblum, H., Gasser, B., Maurer, M., Borth, N. & Mattanovich, D. (2004). Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. *Biotechnol Bioeng* **85**, 367–375.

Jones, J. G. & Bellion, E. (1991). Methanol oxidation and assimilation in *Hansenula polymorpha*. An analysis by ^{13}C n.m.r. *in vivo*. *Biochem J* **280**, 475–481.

Jones, E. W. & Fink, G. R. (1982). Regulation of amino acid and nucleotide biosynthesis in yeast. In *The Molecular Biology of the Yeast Saccharomyces – Metabolism and Gene Expression*, pp. 181–299. Edited by J. N. Strathern, E. W. Jones & J. R. Broach. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Lin Cereghino, G. P., Lin Cereghino, J., Ilgen, C. & Cregg, J. M. (2002). Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Curr Opin Biotechnol* **13**, 329–332.

Lin Cereghino, J. & Cregg, J. M. (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* **24**, 45–66.

Maaheimo, H., Fiaux, J., Çakar, Z. P., Bailey, J. E., Sauer, U. & Szyperski, T. (2001). Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional ^{13}C labeling of common amino acids. *Eur J Biochem* **268**, 2464–2479.

Michal, G. (1998). *Biochemical Pathways: an Atlas of Biochemistry and Molecular Biology*. New York: Wiley.

Minning, S., Serrano, A., Ferrer, P., Solà, C., Schmid, R. D. & Valero, F. (2001). Optimisation of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*. *J Biotechnol* **86**, 59–70.

Prinz, B., Schultchen, J., Rydzewski, R., Holz, C., Boettner, M., Stahl, U. & Lang, C. (2004). Establishing a versatile fermentation and purification procedure for human proteins expressed in the yeasts *Saccharomyces cerevisiae* and *Pichia pastoris* for structural genomics. *J Struct Funct Genomics* **5**, 29–44.

Sauer, U., Hatzimanikatis, V., Bailey, J. E., Hochuli, M., Szyperski, T. & Wüthrich, K. (1997). Metabolic fluxes in riboflavin-producing *Bacillus subtilis*. *Nat Biotechnol* **15**, 448–452.

Sauer, U., Lasko, D. R., Fiaux, J., Hochuli, M., Glaser, R., Szyperski, T., Wüthrich, K. & Bailey, J. E. (1999). Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism. *J Bacteriol* **181**, 6679–6688.

Sauer, M., Branduardi, P., Gasser, B., Valli, M., Maurer, M., Porro, D. & Mattanovich, D. (2004). Differential gene expression in recombinant *Pichia pastoris* analysed by heterologous DNA microarray hybridisation. *Microb Cell Fact* **3**, 17.

Solà, A. (2004). *Estudi del metabolisme central del carboni de Pichia pastoris*. PhD thesis, Universitat Autònoma de Barcelona, Catalonia, Spain.

Solà, A., Maaheimo, H., Ylölen, K., Ferrer, P. & Szyperski, T. (2004). Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*. *Eur J Biochem* **271**, 2462–2470.

Stratton, J., Chiruvolu, V. & Meagher, M. (1998). High cell-density fermentation. *Methods Mol Biol* **103**, 107–120.

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**, 433–448.

Szyperski, T. (1998). ^{13}C-NMR, MS and metabolic flux balancing in biotechnology research. *Q Rev Biophys* **31**, 41–106.

Szyperski, T., Glaser, R. W., Hochuli, M., Fiaux, J., Sauer, U., Bailey, J. & Wüthrich, K. (1999). Bioreaction network topology and metabolic flux ratio analysis by biosynthetic fractional ^{13}C-labeling and two dimensional NMR spectroscopy. *Metab Eng* **1**, 189–197.

Takada, Y. & Noguchi, T. (1985). Characteristics of alanine : glyoxylate aminotransferase from *Saccharomyces cerevisiae*, a regulatory enzyme in the glyoxylate pathway of glycine and serine biosynthesis from tricarboxylic acid-cycle intermediates. *Biochem J* **231**, 157–163.

Voet, D. & Voet, J. G. (1995). *Biochemistry*. New York: Wiley.

Yokoyama, S. (2003). Protein expression systems for structural genomics and proteomics. *Curr Opin Chem Biol* **7**, 39–43.

Zhang, W., Hywood Potter, K. J., Plantz, B. A., Schlegel, V. L., Smith, L. A. & Meagher, M. M. (2003). *Pichia pastoris* fermentation with mixed-feeds of glycerol and methanol: growth kinetics and production improvement. *J Ind Microbiol Biotechnol* **30**, 201–215.

Edited by: M. Tien

PUBLICATION III

^{13}C-metabolic flux ratio and novel carbon path analyses confirmed that *Trichoderma reesei* uses primarily the respirative pathway also on the preferred carbon source glucose

In: BMC Systems Biology 2009. 3:104.

BMC Systems Biology

Research article

^{13}C-metabolic flux ratio and novel carbon path analyses confirmed that *Trichoderma reesei* uses primarily the respirative pathway also on the preferred carbon source glucose

Paula Jouhten*[1], Esa Pitkänen[2], Tiina Pakula[1], Markku Saloheimo[1], Merja Penttilä[1] and Hannu Maaheimo[1]

Address: [1]VTT Technical Research Centre of Finland, Espoo, Finland and [2]Department of Computer Science, University of Helsinki, Helsinki, Finland

Email: Paula Jouhten* - paula.jouhten@vtt.fi; Esa Pitkänen - epitkane@cs.helsinki.fi; Tiina Pakula - tiina.pakula@vtt.fi; Markku Saloheimo - markku.saloheimo@vtt.fi; Merja Penttilä - merja.penttila@vtt.fi; Hannu Maaheimo - hannu.maaheimo@vtt.fi

* Corresponding author

Published: 29 October 2009

BMC Systems Biology 2009, **3**:104 doi:10.1186/1752-0509-3-104

Received: 16 July 2009
Accepted: 29 October 2009

This article is available from: http://www.biomedcentral.com/1752-0509/3/104

Abstract

Background: The filamentous fungus *Trichoderma reesei* is an important host organism for industrial enzyme production. It is adapted to nutrient poor environments where it is capable of producing large amounts of hydrolytic enzymes. In its natural environment *T. reesei* is expected to benefit from high energy yield from utilization of respirative metabolic pathway. However, *T. reesei* lacks metabolic pathway reconstructions and the utilization of the respirative pathway has not been investigated on the level of *in vivo* fluxes.

Results: The biosynthetic pathways of amino acids in *T. reesei* supported by genome-level evidence were reconstructed with computational carbon path analysis. The pathway reconstructions were a prerequisite for analysis of *in vivo* fluxes. The distribution of *in vivo* fluxes in both wild type strain and *cre1*, a key regulator of carbon catabolite repression, deletion strain were quantitatively studied by performing ^{13}C-labeling on both repressive carbon source glucose and non-repressive carbon source sorbitol. In addition, the ^{13}C-labeling on sorbitol was performed both in the presence and absence of sophorose that induces the expression of cellulase genes. Carbon path analyses and the ^{13}C-labeling patterns of proteinogenic amino acids indicated high similarity between biosynthetic pathways of amino acids in *T. reesei* and yeast *Saccharomyces cerevisiae*. In contrast to *S. cerevisiae*, however, mitochondrial rather than cytosolic biosynthesis of Asp was observed under all studied conditions. The relative anaplerotic flux to the TCA cycle was low and thus characteristic to respiratory metabolism in both strains and independent of the carbon source. Only minor differences were observed in the flux distributions of the wild type and *cre1* deletion strain. Furthermore, the induction of the hydrolytic gene expression did not show altered flux distributions and did not affect the relative amino acid requirements or relative anabolic and respirative activities of the TCA cycle.

Conclusion: High similarity between the biosynthetic pathways of amino acids in *T. reesei* and yeast *S. cerevisiae* was concluded. *In vivo* flux distributions confirmed that *T. reesei* uses primarily the respirative pathway also when growing on the repressive carbon source glucose in contrast to *Saccharomyces cerevisiae*, which substantially diminishes the respirative pathway flux under glucose repression.

Background

The industrially important protein producer, the filamentous fungus *Trichoderma reesei*, a clonal derivative of the ascomycete *Hypocrea jecorina*, is adapted to growth in nutrient poor environments, where it is able to use complex plant material as carbon source. *T. reesei* and a number of other filamentous fungi and cellulolytic bacteria produce and secrete plant polymer hydrolyzing enzymes, such as cellulases and hemicellulases, into their surroundings to break down the polymers into easily metabolizable monomers [1].

Because of its ability to synthesize and secrete large amounts of proteins, *T. reesei* has gained industrial importance in production of enzymes of native and heterologous origin. Carbon catabolite repression (CCR) of *T. reesei* negatively regulates the powerful production machinery of the hydrolytic enzymes when a preferred carbon source, such as glucose, is available. Inducers of hydrolytic enzyme expression are often small oligosaccharides or derivative parts of the polymers from the environment of the fungus. The inductive signaling leads to synthesis of specific sets of enzymes [2,3]. In *T. reesei*, D-xylose, xylobiose, sophorose, and lactose have been observed to trigger production of particular enzyme sets [4,5]. Sophorose, a molecule of two beta-1,2-linked glucose units, is an efficient inducer of cellulose gene expression at low concentration (1-2 mM) when *T. reesei* is growing on a non-repressing carbon source, such as sorbitol or glycerol [6]. However, in high glucose concentrations CCR overrules the inductive signals in *T. reesei* [6].

Sorbitol as a carbon source neither provokes CCR nor triggers the cellulase gene expression in *T. reesei* [6]. Nevertheless, cellulase production is positively correlated with the ability of different *T. reesei* strains to grow on D-sorbitol [7], which could be converted to L-sorbose [8] that induces cellulase expression in *T. reesei* [9]. In *T. reesei* L-arabinitol 4-dehydrogenase (Lad1) is involved in the initial oxidization of D-sorbitol at C2 to convert it to D-fructose [10]. A specific sorbitol dehydrogenase converts sorbitol to fructose in *Aspergilli* fungi [11,2].

Cre1 is the key mediator protein of CCR in *T. reesei* [12,13]. *Trichoderma* Cre1 has a 95% sequence similarity with *Aspergillus* CreA in regions of the zinc-finger and proline-serine-threonine-rich domain and the complete sequences are 46% identical [13]. Cre1 is structurally also highly similar to Mig1, a key protein in glucose repression in yeast *Saccharomyces cerevisiae* [12,13]. However, the functional dissimilarities observed between Cre1/CreA and Mig1, in spite of the sequence and structural similarity, have led to the conclusion that glucose repression functionalities in filamentous fungi and yeasts have evolved separately [14,15]. Pfeiffer *et al* argued that the evolution from unicellular to undifferentiated multicellular organisms, like *T. reesei*, has been facilitated by the general preference of high yield energy generation through respiration even in the presence of a preferred carbon source [16]. In contrast to CCR regulation in *S. cerevisiae*, it has been shown that in *T. reesei* CCR does not cause repression of genes encoding the TCA cycle enzymes or respiratory pathway components [17,18]. David *et al* observed differences in the distribution of intracellular carbon fluxes in central carbon metabolism between *A. nidulans* reference and a carbon repression deletion mutant (*creA*Δ4) strains when they were grown on glucose [19].

Despite the industrial importance of *T. reesei*, its genome has only recently been sequenced [20] and its metabolism, beyond that related to protein production and secretion, is narrowly studied. In the present work, computational carbon path analysis [21] was utilized to reconstruct the biosynthetic pathways of amino acids. That was essential for quantitative flux analysis, as no metabolic network model was available for *T. reesei*. The localizations of the key reactions in the biosynthetic pathways of amino acids were determined from the ^{13}C-labeling patterns of proteinogenic amino acids and by computational estimation of targeting peptide sequences. The intracellular metabolic flux ratios in the central carbon metabolism were determined utilizing fractional ^{13}C-labeling and metabolic flux ratio (METAFoR) analysis [22] in a wild type (QM6a) strain and in a Δ*cre1* mutant strain (L161a), when grown on the repressive carbon source glucose and on the neutral carbon source sorbitol. Additionally, the effect of sophorose induction of cellulase gene expression on the relative fluxes in the central carbon metabolism was quantified. To the authors' knowledge this is the first time that the metabolic pathways of *T. reesei* have been reconstructed and *in vivo* fluxes in the central carbon metabolism of *T. reesei* have been quantitatively studied.

Results and Discussion

^{13}C-labeling in batch cultures

Metabolic flux ratio (METAFoR) analysis was performed for the *T. reesei* wild type (QM6a) and Δ*cre1* (L161a) strains growing in minimal medium in flasks with fractional [U-^{13}C]glucose and on fractional [U-^{13}C]sorbitol with and without induction of cellulase gene expression by sophorose. Since ^{13}C-metabolic flux ratio (METAFoR) analysis is based on biosynthetically directed fractional (BDF) labeling of the constituents of biomass biopolymers, it requires constant intracellular flux distribution during the labeling [22-28]. Constant flux distribution can be achieved in a chemostat culture, where the specific growth rate is constant, or in a batch cultivation during exponential growth. In the exponential growth phase

when the cells are growing at their maximum specific growth rate and the changes in the extracellular conditions are still insignificant a quasi-steady state can be assumed [26,19]. Precultivations were performed to determine the exponential growth phases and the maximum specific growth rates for the *T. reesei* wild type and Δ*cre1* strains (data not shown). Cultures with different growth profiles were then sampled for quantitative flux analysis at equivalent growth stages, in the exponential phase. The maximum specific growth rates of *T. reesei* on glucose were 0.15 ± 0.01 h^{-1} and 0.12 ± 0.01 h^{-1} for the wild type and Δ*cre1* strains, respectively. When grown on sorbitol the maximum specific growth rates for the wild type and Δ*cre1* strains were 0.03 ± 0.02 h^{-1} and 0.06 ± 0.01 h^{-1}, respectively. The maximum specific growth rates of *A. nidulans* wild type strain and that of a CreA deletion strain have been observed to be 0.25 h^{-1} and 0.11 h^{-1}, respectively, when grown on glucose [19].

Reconstruction of pathways through the metabolic network leading to amino acid synthesis

In order to quantify the in vivo flux ratios in the central carbon metabolism of *T. reesei* by ^{13}C-labelling and METAFoR approach [22], it was necessary to obtain a model of the amino acid biosynthesis pathways. However, no curated metabolic model exists for *T. reesei*. Thus, the pathways for synthesis of the carbon backbones of the proteinogenic amino acids from the carbon source molecules in *T. reesei* were reconstructed by *ReTrace* pathway analysis [21]. ReTrace analysis results are summarized in Table 1 and fully reported in Additional File 1.

ReTrace analysis confirmed, for most of the proteinogenic amino acids, that the biosynthetic pathways of amino acids identical to the pathways in *S. cerevisiae* are present also in *T. reesei*. Therefore the carbon backbones of the proteinogenic amino acids in *T. reesei* evidently originate from the precursor metabolites similar to the ones in *S. cerevisiae* [26] (Figure 1). For some proteinogenic amino acids (Arg, Ile, Leu, Thr, Tyr) ReTrace was not directly able

Table 1: Summary of the ReTrace [21] analysis of the amino acid biosynthetic routes in *T. reesei*.

Amino acid	Precursors	Paths	Zo	AvgSc	BestSize	AvgSize	MinPoor
Ala	Pyr	227	1	652	1	16.8	0
Arg	Oga	134	1	811	9	15.1	0
Asp	Oaa	121	1	757	1	13.5	0
Glu	Oga	36	1	426	1	12.6	0
Gly	Ser	260	1	1128	1	16.1	0
Gly	Thr	71	1	522	2	12.6	0
His	R5P	21	1	774	22	25.8	0
Ile	OAA, Pyr	483	1	797	14	18.7	0
Leu	AcCoA, Pyr	916	1	356	13	19.1	1
Lys	AcCoA, Oga	347	0.67	834	11	14.6	0
Phe	E4P, PEP	348	1	679	12	19.2	0
Pro	Oga	119	1	673	3	14.4	0
Ser	3PG	69	1	670	3	15.3	0
Thr	Oaa	97	1	768	7	2.5	0
Tyr	E4P, PEP	156	1	654	19	19.6	0

Paths paths found, Zo highest fraction of transferred atoms, AvgSc average reaction scores, BestSize the size of the pathway achieving Zo reported, AvgSize average pathway size, MinPoor minimum number of reactions with a score under 50

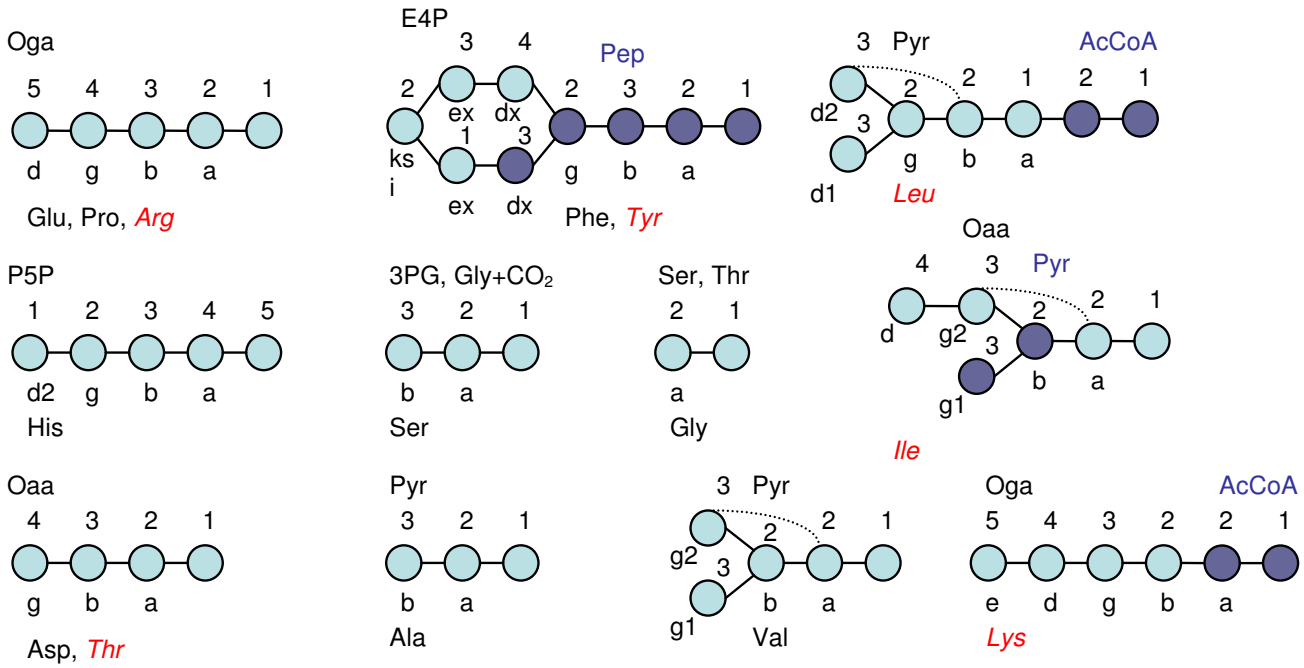

Figure 1

Origins of proteinogenic amino acids. The origins of the carbon backbones of the proteinogenic amino acids utilized in METAFoR analysis [26] and for which the biosynthetic pathways were reconstructed by computational pathway analysis method ReTrace [21]. The amino acids for which the biosynthetic pathway was not directly found by ReTrace are denoted in red italics. The amino acid carbons are denoted in the following way: a = α, b = β, g = γ, d = δ, e = ε, ksi = ξ.

to identify the biosynthetic routes that are active in *S. cerevisiae* because alternative reactions with higher scores strongly directed the search or because of errors in the atom mapping in the KEGG reaction database. However, the manual inspection of all the pathways identified directly from the carbon source or from different precursors, confirmed that the biosynthetic pathways for all proteinogenic amino acids relevant for METAFoR analysis, except for Arg and Lys, that are known to operate in *S. cerevisiae*, are also present in *T. reesei*.

The fungal biosynthetic pathway of Lys from Oga [29] was not found by ReTrace because of inconsistencies in the atom mapping in the KEGG reaction database. However, because the reactions of the alternative biosynthetic route of Lys which is active, for example in bacteria, did not gain good scores for presence in *T. reesei*, the fungal pathway was assumed prior to the ^{13}C-pathway analysis. Pathways from Oga to Arg were identified by ReTrace but the pathway known to be active in *S. cerevisiae* was not found among them. Most of the identified pathways were directed through 1-pyrroline-5-carboxylate dehydrogenase reaction (1.5.1.12) in the reverse direction, which forms a false shortcut path between Oga and Arg.

After the unsuccessful direct search of pathway from Oaa to Thr, Thr biosynthesis pathway was searched from Asp, an intermediate in the pathway from Oaa to Thr. Genome level evidence of the presence of the pathway was found. The biosynthetic pathway of Ile that is active in *S. cerevisiae* was not found directly from precursors Oaa and Pyr because the pathway proceeds first from Oaa to Thr and that pathway was not directly identified as discussed above. The reactions further from Thr were identified with high scores for genome level evidence of their presence in *T. reesei* and thus, the pathway that is known to be active in *S. cerevisiae* is evidently present also in *T. reesei*. Tyr biosynthesis pathway was found from precursors downstream to 3-(4-hydroxyphenyl)pyruvate and only the transamination was lacking from a complete pathway. However, a high scoring hit for a transaminase sequence was separately searched and identified in the genome of *T. reesei*. Most of the high scoring alternative pathways could be excluded because only the anabolic pathways are active in the exponential growth and in absence of amino acids in the medium.

^{13}C-pathway analysis and prediction of subcellular localization of key enzymes

The pathways of amino acid biosynthesis reconstructed in *T. reesei* corresponded to the pathways utilized by *S. cerevisiae*. The fragmentomer data from ^{13}C-labeling of proteinogenic amino acids provided further confirmation for this (see Methods for the definition of fragmentomer data). The ^{13}C-labeling patterns of the carbon backbones of proteinogenic amino acids originate from the ^{13}C-labeling of their precursor metabolites in central carbon metabolism and thus, the ^{13}C-labeling patterns of amino acids can be propagated to the precursors to identify the active pathways. In particular, the Lys ^{13}C-labeling pattern indicated its synthesis from Oga via α-aminoadipate pathway, as in yeasts [30]. However, in contrast to *S. cerevisiae* and a number of other yeast [26,31], the ^{13}C-labeling pattern of Asp indicated that it primarily originated from mitochondrial Oaa under all the studied conditions (Figure 2). Mitochondrial Asp synthesis has previously been observed in *Yarrowia lipolytica* [31]. Furthermore, identical ^{13}C-labeling patterns were observed in Asp and Thr. This confirmed Thr synthesis from Asp and excluded a contribution from the reversible threonine aldolase reaction [32].

A three carbon fragment of mitochondrial Oaa (Oaa_{mit}) (C2-C3-C4) remains intact in the synthesis of the TCA cycle intermediate Oga and therefore the ^{13}C-labeling pattern of Oaa_{mit} can be partially observed in Glu that originates from Oga. In the exponential growth phase it is reasonable to assume unidirectional transport of Oaa across the mitochondrial membrane, which has previously been experimentally shown in *S. cerevisiae* [26]. When the backward transport from Oaa_{mit} to Oaa_{cyt} is negligible, a three-carbon fragment of Oaa_{cyt} (C1-C2-C3) is produced from Pep, a precursor of Phe and Tyr, via glycolysis and by pyruvate carboxylase. The fractions of intact Cα-Cβ bonds in Asp and Thr were highly similar to the corresponding intact carbon fragments in Oga, propagated from Glu, but clearly different from the corresponding intact carbon fragments in Pep, propagated from Phe, Tyr in *T. reesei*, indicating the primarily mitochondrial origin of Asp. Since the C3-C4 bond of Oaa_{mit} remains intact in the TCA cycle, Oaa_{mit} serving as a precursor for Asp and Thr biosynthesis was further supported by the high similarity in the fraction of molecules having the corresponding C-C fragment intact, i.e. the C3-C4 fragment of Oaa, propagated from Asp, Thr, and C1-C2 fragment of Oga, propagated from Glu, (OAA_xx1 and OGA_1xxx, respectively, Figure 2).

Additional support for mitochondrial Asp synthesis was obtained from sequence analysis. Evidence of mitochondrial targeting peptide sequence was identified in one of the *T. reesei* genome sequences with homology to the aspartate aminotransferases in *S. cerevisiae* by TargetP analysis [33,34] (Additional file 2). This strongly supported the mitochondrial localization of one of the encoded enzymes. However, no evidence of mitochondrial targeting peptide was identified by TargetP analysis of the *T. reesei* sequence with homology to the *S. cerevisiae* pyruvate carboxylase that produces Oaa. Thus, Oaa_{mit} could originate both from transport across the mitochondrial membrane and from the TCA cycle.

Pyruvate is a precursor of Ala and Val biosynthesis. If the pyruvate pools in cytosol and mitochondria possess significantly different ^{13}C-labeling patterns, for example as a result of malic enzyme flux, a mitochondrial localization of pyruvate-based amino acid synthesis can be confirmed from the ^{13}C-labeling data [31]. However, the fractions of intact two-carbon fragments Pyr C1-C2 and C2-C3 observed in Ala and Val and the corresponding two-carbon fragments in Pep, a direct precursor of Pyr_{cyt}, observed in Phe and Tyr, were not significantly different under the studied conditions. Therefore, the ^{13}C-labeling patterns could not be utilized to assess the localization of the synthesis of pyruvate-based amino acids. Strong evidence of a mitochondrial targeting sequence in the *T. reesei* sequence that showed homology to the acetolactate synthase in *S. cerevisiae* was identified by TargetP [33,34] (Additional file 2). In yeast the first enzyme in Val biosynthesis, the acetolactate synthase, has been reported to be localized in mitochondria [35], whereas cytosolic and mitochondrial isoenzymes of alanine aminotransferase have been observed [30].

Ser originates from glycolytic intermediate 3-phosphoglycerate and can be further converted to Gly and a C1 unit by the reversible reaction of serine hydroxymethyl transferase (SHMT). Gly could also originate from threonine aldolase or from the reversible reaction of the glycine cleavage pathway (i.e., C1 + CO_2). In *S. cerevisiae* glycine cleavage pathway is active inside mitochondria [36] and although both mitochondrial and cytosolic isoenzymes of SHMT exist in *S. cerevisiae* [37,38], the effect of the glycine cleavage pathway on the Ser-Cα *f(1)* fraction has not been observed in *S. cerevisiae* batch cultures grown on glucose [26]. In *T. reesei* the activity of the glycine cleavage pathway was observed in the ^{13}C-labeling pattern of Ser, since Ser-Cα *f(1)* fragmentomer fraction of Ser molecules with both carbon bonds cleaved was higher than the fraction of fully cleaved Pep, a three carbon lower glycolytic intermediate. The fraction of fully cleaved Pep molecules was observed in Phe and Tyr-Cα *f(1)* fragmentomer fractions (Figure 3). Two *T. reesei* sequences were observed to have homology to the *S. cerevisiae* SHMT sequences. In one of them a strong evidence of a mitochondrial targeting presequence was found by TargetP [33,34] (Additional file 2). Therefore, SHMT activity likely occurs in both

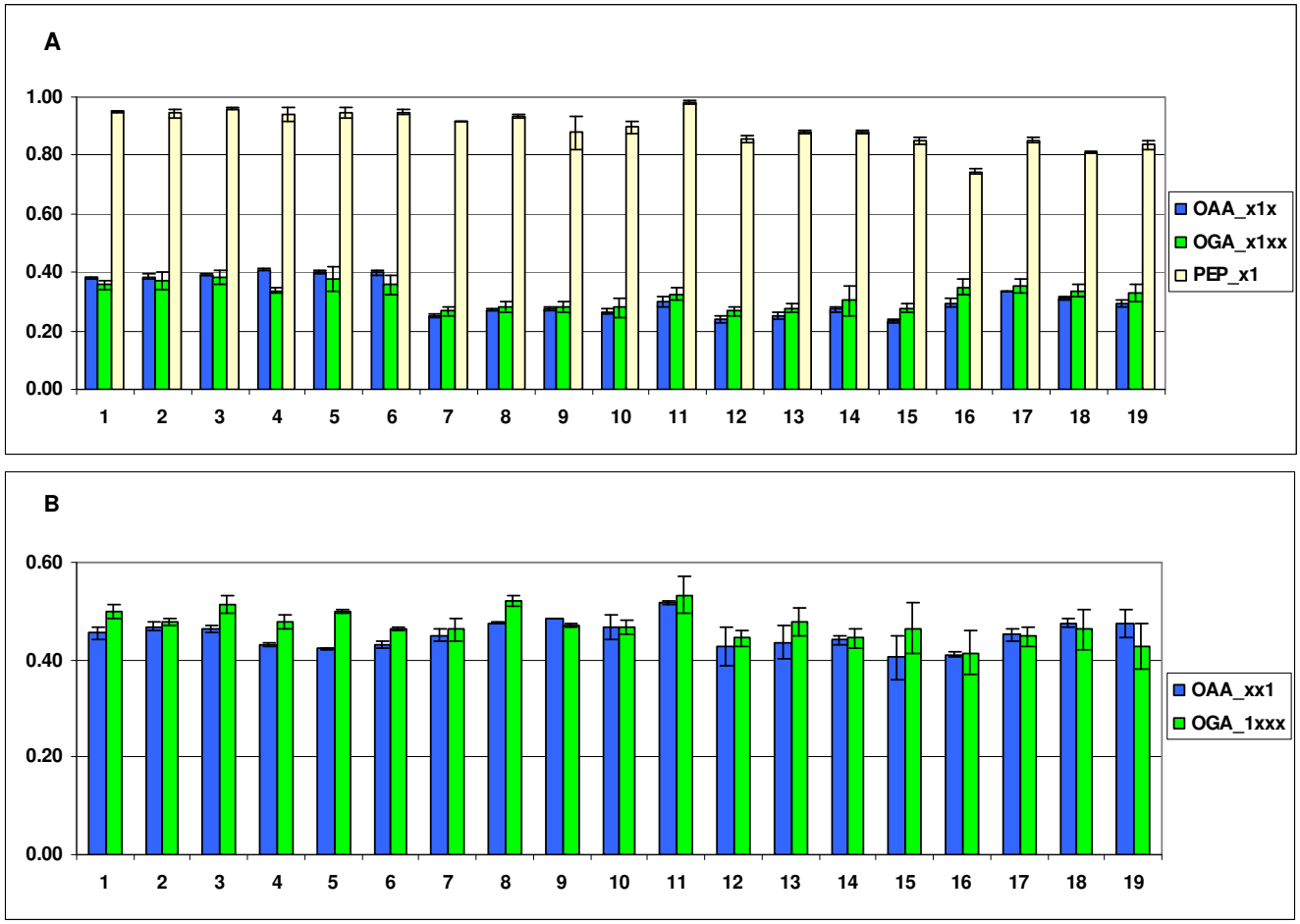

Figure 2

Comparison of the fractions of corresponding intact bonds in amino acid precursors Oaa, Oga and Pep in *T. reesei*. The data is taken from all replicates of fractional [U-^{13}C]glucose or [U-^{13}C]sorbitol experiments performed (1-3 wild type on glucose, 4-6 Δ*cre1* on glucose, 7-9 wild type on sorbitol, 10-12 wild type on sorbitol (sophorose experiment control), 13-15 wild type on sorbitol (sophorose induction), 16-17 Δ*cre1* on sorbitol (sophorose experiment control), 18-19 Δ*cre1* on sorbitol (sophorose induction)). Oaa data was detected from Asp-Cα, -Cβ and Thr-Cα, Oga data from Glu-Cα, -Cβ and Pro-Cα, -Cβ and Pep data from Phe and Tyr-Cα. Fractions of intact bonds in Oaa, Oga and Pep were calculated from combinations of fragmentomers. A) OAA_x1x is the fraction of Oaa molecules with an intact bond at C2-C3, OGA_x1xx is the fraction of Oga molecules with an intact bond at C2-C3 and PEP_x1 is the fraction of Pep molecules with an intact bond at C2-C3. B) OAA_xx1 is the fraction of Oaa molecules with an intact bond at C3-C4 and OGA_1xxx is the fraction of Oga molecules with an intact bond at C1-C2. Error bars are ± SEMs. The carbon chain of Oaa_{mit} remains intact in the TCA cycle except that C1 is cleaved in the synthesis of Oga. Almost the entire labeling pattern of Oaa_{mit} can be assessed from the labeling pattern determined for Oga. If Asp and Thr synthesis originates from Oaa_{mit}, the fractions of corresponding Oaa and Oga intact fragments in the figures should match.

cytosolic and mitochondrial compartments of *T. reesei*. The Ser ^{13}C-labeling pattern observed in *T. reesei* indicates either a partially cytosolic localization of the glycine cleavage pathway or protein synthesis occurring primarily from a mitochondrial pool of Ser. TargetP analysis of the *T. reesei* sequence homological to sequence of the *S. cerevisiae* glycine dehydrogenase, the p-subunit of the Gly cleavage system, showed no clear indication of a mitochondrial targeting pre-sequence [33,34] (Additional file 2).

Amino acids belonging to the Glu amino acid family, Glu, Pro and Arg, showed a highly similar ^{13}C-labeling, as expected, in both strains grown on glucose and in the wild type strain culture grown on sorbitol (data not shown). In contrast, a significant variation was observed in the Glu,

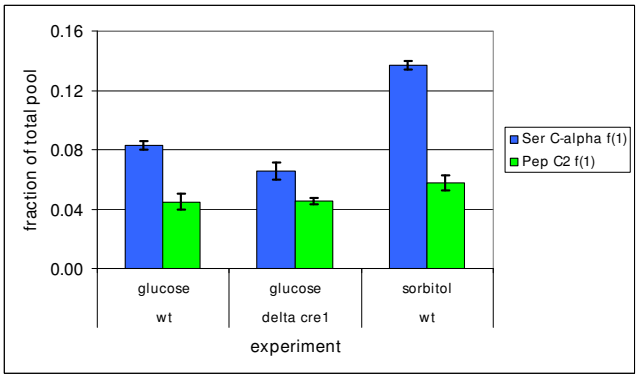

Figure 3
Effect of the reversible glycine cleavage pathway. Effect of the reversible glycine cleavage pathway in *T. reesei* wild type (wt) and Δ*cre1* strains on the ^{13}C-labeling pattern of Ser. The fraction of Ser-Cα *f(1)* fragmentomer from the total pool of Ser, compared to the corresponding fraction in Pep (C2) observed from Phe and Tyr-Cα *f(1)* fragmentomers. Error bars are ± SEMs.

Pro and Arg fragmentomers in cultures when their pre-cultures were mixed before sophorose induction experiment. This observation may be explained by differential mobilization of amino acids from cellular compartments resulting from the perturbation when the cultures were mixed prior to the induction period.

METAFoR analysis of T. reesei

The relative fluxes merging at the branching points of central carbon metabolism of *T. reesei* were determined by BDF ^{13}C-labeling with glucose and sugar alcohol sorbitol as carbon sources. The flux ratios of the wild type and Δ*cre1 T. reesei* strains determined in batch cultures on glucose, on sorbitol and on sorbitol with sophorose induction of expression of cellulase genes are shown in Table 2.

Respiratory pathway flux of T. reesei

The relative anaplerotic flux (the fraction of OAA_{mit} from Pep, Table 2) describes the relative activities of the biosynthetic and respirative carbon fluxes in the TCA cycle. The anaplerotic flux replenishes carbons to the TCA cycle flux by importing C4 compounds as there is drain of carbon to biosynthesis. David *et al* concluded that the TCA cycle was more active in Δ*creA* mutants of *A. nidulans* than in the wild type when grown on glucose [19]. However, they suggested that it resulted from a higher ATP demand in the deletion strain possibly caused by active futile cycles instead of derepression of the respirative pathway flux. No significant difference was observed in the anaplerotic flux ratios between the *T. reesei* wild type and Δ*cre1* strains grown on glucose. Thus, in *T. reesei* the TCA cycle was as active relative to biosynthesis in the Δ*cre1* strain as in the wild type strain and so, Cre1 does not mediate repression of respirative pathway flux in *T. reesei* either.

A difference was observed between the two *T. reesei* strains when grown on sorbitol. The relative anaplerotic flux was

Table 2: Metabolic flux ratios of *T. reesei* wild type (wt) and Δ*cre1* strains in aerobic batch cultures on glucose and on sorbitol with and without sophorose induction of cellulase gene expression.

strain	wt		Δ*cre1*		wt		wt		Δ*cre1*		Δ*cre1*	
carbon source	glucose		glucose		sorbitol		sorbitol		sorbitol		sorbitol	
fraction of total pool (%)		sd		sd		sd	sophorose	sd	control	sd	sophorose	sd
Pep from PPP (UB wo PEPck)	39	2	47	4	36	7	37	7	45	9	46	2
R5P from T3P and S7P	51	1	42	1	72	3	70	1	79	4	79	1
R5P from E4P	25	2	23	1	46	2	48	5	54	3	44	2
Ser from Gly and C1	80	0	85	2	44	2	41	15	54	1	52	0
Gly from CO_2 and C1	12	1	14	1	6	2	16	21	8	1	10	3
Oaa_{mit} from Pep	35	1	33	2	26	3	26	4	42	5	39	7
MAE (UB)	4	0	9	1	12	2	2	2	6	5	1	nd
MAE (LB)	2	0	6	1	9	1	2	2	4	3	0	nd

sd standard deviation, UB upper bound, LB lower bound

26% in the wild type strain and 42% in the Δ*cre1* strain (Table 2). This may indicate that there was a difference in the specific growth rates of the two strains on sorbitol.

Previously, an excess of glucose has been found to only partially repress the gene expression of the enzymes of the TCA cycle and the components of the respiratory chain in *T. reesei* [17]. That is in contrast to the effect of excess of glucose on *S. cerevisiae*, where glucose repression extensively downregulates the respiratory pathway at the transcriptional level [18]. The anaplerotic flux ratio in *T. reesei* wild type strain was higher on glucose (35%) than that on sorbitol (26%) (Table 2). The results indicated a higher activity of respiratory metabolism relative to biosynthesis on the non-repressing carbon source sorbitol than that on the repressing carbon source glucose. A complete oxidation of sorbitol, that is a more reduced carbon source than glucose, results in a higher relative flux of electrons per carbon source molecule to the respiratory chain than during growth on glucose. Thus, if *T. reesei* respired at maximum rate during the batch growth on glucose, fluxes producing reduced cofactors, for example biosynthetic pathway fluxes or the TCA cycle fluxes, would have decreased on sorbitol.

Small fractions of Pyr_{mit} originating from malate via the action of the malic enzyme were observed in both strains under almost all conditions (Table 2).

Pentose phosphate pathway (PPP) of T. reesei

A lower fraction of triose phosphates originated from pentose phosphates in the wild type strain (39%) than in the Δ*cre1* strain (47%) when grown on glucose (Table 2). In batch cultures under excess glucose conditions, the gluconeogenesis by phosphoenolpyruvate carboxykinase and the reverse transport of Oaa across the mitochondrial membrane are assumed to have negligible fluxes [18,26]. For this purpose the fraction of Pep originating from the pentose phosphate pathway (PPP) was calculated neglecting any contribution of phosphoenolpyruvate carboxykinase to the ^{13}C-labeling pattern of Pep, propagated from Phe, Tyr. The fraction of Pep originating from PPP represents the flux via PPP relative to the total flux to Pep. However, this fraction is not a direct measure of the flux through the oxidative branch of the PPP but includes molecules that have only gone through reversible reactions in the non-oxidative PPP. Furthermore, the standard deviation is always high because only 40% of the triose phosphates that originate from the PPP have different ^{13}C-labeling patterns than the triose phosphates originating from glycolysis.

The differences in the relative flux through the PPP to the triose phosphates can be caused by differences in the glycolytic rate or in NADPH demands, since the oxidative branch of the PPP is usually the main source of cytosolic NADPH. A low glycolytic rate could allow the label scrambling in the non-oxidative part of the PPP to affect the ^{13}C-labeling patterns of a large fraction of triose phosphates.

The reversible fluxes through the reactions of transketolase and transaldolase, observed in the ^{13}C-labeling patterns of pentose phosphates that can be detected in His, were clearly different in glucose and sorbitol cultivations (Table 2). The fraction of pentose phosphates that had gone through a transketolase reaction (R5P from T3P and S7P) was 51% and 42% when glucose was the carbon source for the wild type and the Δ*cre1* strains, respectively. When grown on sorbitol the fractions were higher, 72% and 79% for the wild type and the Δ*cre1* strains, respectively. The fraction of pentose phosphates cleaved in the transaldolase and transketolase reactions (R5P from E4P) was 25% and 23% when grown on glucose, whereas when sorbitol was the carbon source they were 46% and 54% for the wild type and the Δ*cre1* strains, respectively. The higher fractions of pentose phosphates cleaved in the reactions of transketolase or transaldolase when grown on sorbitol could be a result of entrance of sorbitol in the central carbon metabolism and into the PPP directly in a form of fructose 6-phosphate [10].

Figure 4 shows the relative abundances of the contiguous ^{13}C-fragments around His-Cβ, which originate from fragments around ribose 5-phosphate C3. When sorbitol was the carbon source lower fractions of fully intact His fragments and higher fractions of His fragments cleaved in the reversible reactions of transaldolase and transketolase were observed in both strains than when grown on glucose. This indicated higher relative fluxes in the non-oxidative part of the PPP when compared to the rate of withdrawal of pentose phosphates to His biosynthesis. When sorbitol was the carbon source the relative activity of the non-oxidative PPP compared to the rate of biosynthetic drain of pentose phosphates was even higher in the Δ*cre1* strain than in the wild type strain. The fraction of fully cleaved His-Cβ $f(1)$ fragments was higher in the Δ*cre1* strain than in the wild type strain when they were grown on sorbitol (Figure 4). Correspondingly, lower fractions of fully intact His-Cβ $f(3)$ were observed in the Δ*cre1* strain than in the wild type strain.

Effect of sophorose induction of cellulase gene expression on metabolic fluxes

Induction of cellulase gene expression with sophorose did not cause any significant changes in the metabolic flux distributions in the central carbon metabolism of *T. reesei*. Therefore, the induction of cellulase gene expression did not affect the relative fluxes to different amino acid families or the ratio of anabolic and catabolic activity of the central carbon metabolism. Alteration in the relative bio-

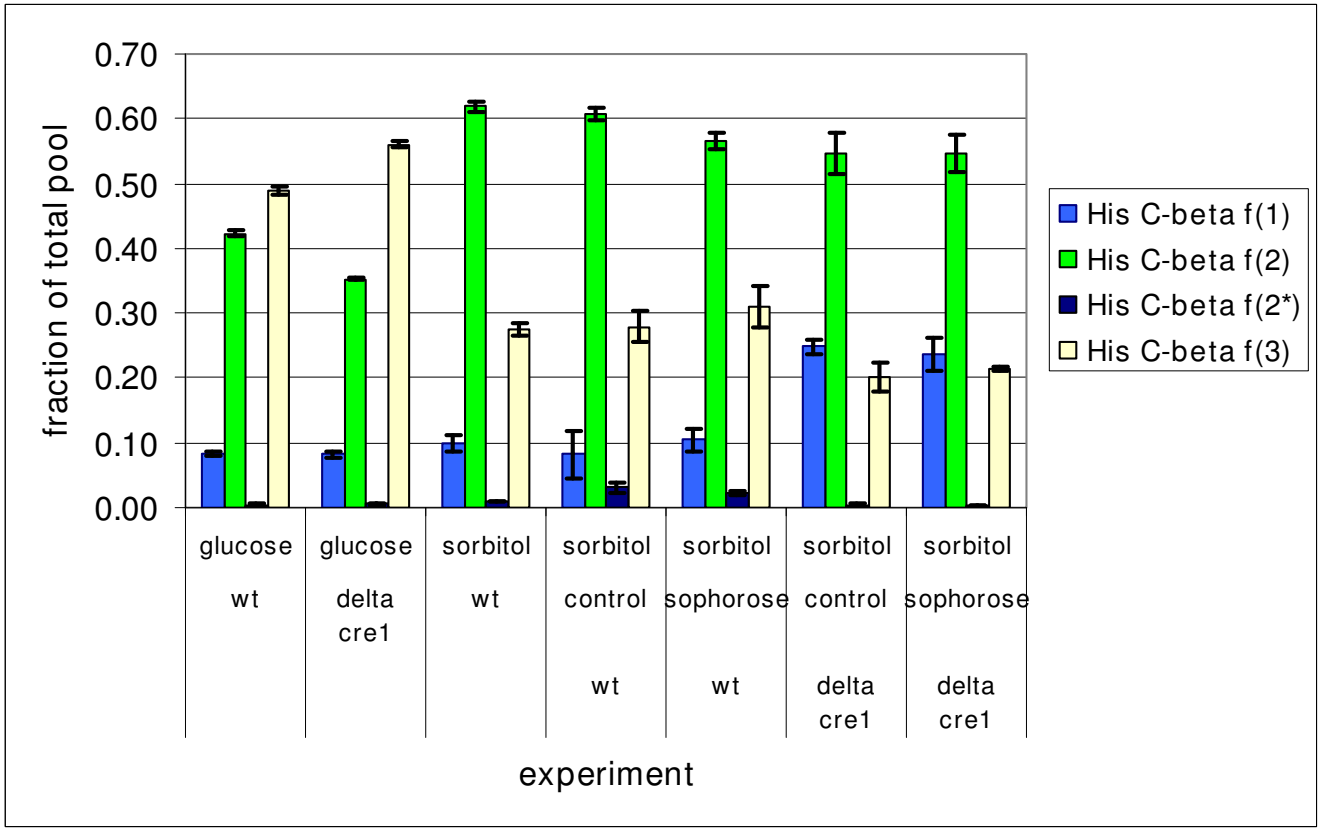

Figure 4
^{13}C-His-Cβ centered contiguous ^{13}C-fragments in *T. reesei*. Fractions of ^{13}C-His-Cβ centered contiguous ^{13}C-fragments in *T. reesei* in different genetic, wild type (wt) and Δ*cre1* mutant. Cultures were grown on glucose (glucose repressed) or sorbitol (derepressed). Sorbitol grown cultures were grown with or without induction of cellulase gene expression by the addition of sophorose to some cultures. His-Cβ *f*(1) denotes fragments with C-C bonds cleaved on both sides of Cβ, His-Cβ *f*(2) and *f*(2*) denote fragments with Cβ-Cα and Cβ-Cγ preserved, respectively, and His- Cβ *f*(3) denotes fragments were both bonds are intact. Error bars represent ± SEM.

synthetic rates of different amino acids would have occurred if the amino acid composition of the induced cellulases had been significantly different from the amino acid composition of the proteins generally produced by *T. reesei* which was not observed.

Flux ratio profiles of T. reesei, S. cerevisiae *and* Pichia stipitis *indicate differences in preferred utilization of pathways*

The anaplerotic flux ratios determined hereby in the wild type and Δ*cre1 T. reesei* strains in batch cultures, both grown on glucose, were substantially lower and similar to what has been observed in fully respiratory metabolism in *S. cerevisiae* in glucose-limited chemostat cultures, where there is no glucose repression [27] (Table 3). The extensive glucose repression of the TCA cycle and the respiratory pathway activity in *S. cerevisiae* result in high anaplerotic ratio in batch cultures on glucose [26]. The anaplerotic flux ratios in the *T. reesei* strains with glucose as a carbon source were also similar to the ones observed in *P. stipitis*, both when grown on glucose in batch cultures and in glucose-limited chemostat cultures [27]. *P. stipitis* completely lacks aerobic alcoholic fermentation.

It has previously been determined that glucose does not cause extensive repression of the gene expression of the TCA cycle and the respiratory pathway components in *T. reesei* [17] as it does in *S. cerevisiae* [18]. The ^{13}C-labeling and METAFoR analysis results on the level of *in vivo* fluxes confirmed that for highly efficient energy generation through complete oxidation of carbon source *T. reesei* indeed uses primarily the respirative pathway also when growing on a preferred carbon source glucose. The regulatory differences between *T. reesei* and *S. cerevisiae* have been explained as adaptation to different growth environments. *S. cerevisiae* is adapted to nutrient rich environments in which it has competitive advantage from fast nutrient utilization and a high rate of ATP production

Table 3: Metabolic flux ratios of *T. reesei* wild type (wt) and Δ*cre1* strains in compared with the corresponding flux ratios observed in the crabtree positive and negative yeasts *S. cerevisiae* and *P. stipitis* [26,27].

organism	***T. reesei***		***T. reesei***		***T. reesei***		***T. reesei***		***S. cerevisiae***		***S. cerevisiae***		***P.stipitis***		***P.stipitis***	
strain	**wt**		**Δ*cre1***		**wt**		**Δ*cre1***									
carbon source	glucose		glucose		sorbitol		sorbitol		glucose		glucose		glucose		glucose	
culture	batch		batch		batch		batch		batch		chemostat		batch		chemostat	
reference							control		[26]		[27]		[27]		[27]	
fraction of total pool (%)		sd		sd		sd		sd		sd		sd		sd		sd
Pep from PPP (UB wo PEPck)	39	2	47	4	36	7	45	9	0-4		40	8	57	9	61	11
R5P from T3P and S7P	51	1	42	1	72	3	79	4	68	2	59	2	57	2	72	2
R5P from E4P	25	2	23	1	46	2	54	3	10	2	33	2	35	2	43	2
Oaa_{mit} from Pep	35	1	33	2	26	3	42	5	76	4	31	2	36	2	32	2
MAE (UB)	4	0	9	1	12	2	6	5	25-30		<13		<6		<7	
MAE (LB)	2	0	6	1	9	1	4	3	nd	nd	nd	nd	nd	nd	nd	nd

through the fermentative pathway, whereas *T. reesei* is adapted to nutrient poor environments where it benefits from high energy yield [17,16]. It has also been postulated that undifferentiated multicellular organisms, of which *T. reesei* is an example, have gotten evolutionary advantage from preferring the high energy yield from respiratory metabolism [16].

Conclusion

Biosynthetic pathways of *T. reesei* were reconstructed for most of the proteinogenic amino acids by using a computational carbon path analysis method ReTrace. The method was used to search for pathways from a metabolic network consisting of all reactions found in a comprehensive metabolic reaction database, and to subsequently rank the pathways according to the degree of support from the *T. reesei*'s genome [21]. Contiguous pathways, identical to the amino acid biosynthetic routes of *S. cerevisiae*, were found with high genome-level evidence. The ^{13}C-labeling patterns observed in this study were in good accordance with the compartmentalized model of eukaryotic central carbon metabolism, originally developed for *S. cerevisiae* [26]. However, in contrast to *S. cerevisiae*, Asp synthesis was observed to occur primarily from the mitochondrial pool of Oaa in both *T. reesei* strains under all the studied conditions.

The *T. reesei* wild type strain is known to exhibit carbon catabolite repression of hydrolytic gene expression during growth on glucose, whereas in the Δ*cre1* strain the repression is partially disturbed [13]. The respirative pathway in *T. reesei* does not become transcriptionally downregulated by the carbon catabolite repression as in *S. cerevisiae* [17]. However, it is the *in vivo* fluxes that are the ultimate phenotype of an organism. In the present work, the effect of carbon catabolite repression on *in vivo* fluxes in *T. reesei* was, for the first time, quantitatively studied. The relative anaplerotic flux to the respirative pathway flux was characteristic to primarily respiratory metabolism in the both *T. reesei* strains under all studied conditions. Thus, *T. reesei*

utilizes primarily respiratory metabolism also when growing on a preferred carbon source glucose. However, the observed relative anaplerotic fluxes suggested that the respirative activity of the TCA cycle is even slightly higher when *T. reesei* grows on the neutral carbon source sorbitol than when it grows on glucose. Only minor differences were observed between the *in vivo* flux distributions of the wild type and the Δ*cre1 T. reesei* strains. This indicates, that Cre1, the key repressor of utilization of alternative carbon sources, does not mediate carbon source dependent metabolic state alterations in the central carbon metabolism of *T. reesei*. The induction of cellulase gene expression with sophorose did not result in significant changes in the relative requirements of proteinogenic amino acids or in the ratio of anabolic and oxidative activities of the TCA cycle.

Methods

Strains, media and culture conditions

Biosynthetically directed fractional (BDF) ^{13}C-labeling of proteins was carried out for the *T. reesei* QM6a (wild type) [39] and *T. reesei* QM6a with deleted *cre1* gene (unpublished). Both strains were cultivated in triplicate on two different carbon sources: glucose and sorbitol. Glucose cultivations were carried out in 2 l flasks in 200 ml of minimal medium ($(NH_4)_2SO_4$ 7.6 g/l, KH_2PO_4 15.0 g/l, 2.4 mM $MgSO_4$, 4.1 mM $CaCl_2$, $CoCl_2$ 3.7 mg/l, $FeSO_4 \cdot 7H_2O$ 5 mg/l, $ZnSO_4 \cdot 7H_2O$ 1.4 mg/l, $MnSO_4 \cdot 7H_2O$ 1.6 mg/l, pH adjusted to 4.8 with KOH) supplemented with 2% (w/v) glucose containing 10% (w/w) [U-^{13}C]glucose.

The 200 ml cultures were inoculated with 8 × 10^7 spores and cultivated at +28°C with constant agitation at 250 rpm. After 35 h of cultivation, during the exponential growth phase (Figure 5), 30 ml and 50 ml samples were withdrawn for dry weight determination and for NMR experiments, respectively. Mycelium from the samples was collected by filtration through Whatmann GF/B filters and washed twice with the sampling volume of water. For dry weight determination the mycelium was dried in an oven at +106°C overnight and weighed.

The BDF ^{13}C-labeling of the wild type strain on sorbitol was carried out in three replicates with 2% (w/v) sorbitol containing10% (w/w) [U-^{13}C]sorbitol, similarly as in the glucose cultivations. After 104 h of incubation, in the exponential growth phase (Figure 5), 30 ml and 50 ml samples were withdrawn for dry weight determination and for NMR experiments, respectively.

BDF ^{13}C-labeling on sorbitol was also carried out with induction of cellulase gene expression by sophorose. Six 2 l flasks of each strain were inoculated, with 2% (w/v) sorbitol as the sole carbon source in minimal medium (see above). After 76 h for the wild type and after 114 h for the Δ*cre1* mutant, in exponential growth phase (Figure 5), the six cultures were combined, a 30 ml sample for dry weight determination was withdrawn and then the culture broth was redivided into six flasks. The final concentration of 1 mM sophorose was introduced into three of the six replicate 2 l flasks to induce cellulase gene expression. An identical volume of water was added to the three control cultures. Three hours after the induction, when cellulase gene expression was expected to be at a moderate level [6], 0.4 g of [U-^{13}C]sorbitol was added to all six cultures to initiate BDF ^{13}C-labeling. The addition of 0.4 g of [U-^{13}C]sorbitol at this time was estimated to result in a [U-^{13}C]sorbitol fraction of about 10% of the total sorbitol in the culture medium. After 24 h from the addition of the [U-^{13}C]sorbitol, still during the early-exponential growth phase (Figure 5), 30 ml and 50 ml samples were withdrawn for dry weight determination and NMR experiments, respectively. Thereby the ^{13}C-labeled fraction of biomass was synthesized in the induced conditions and

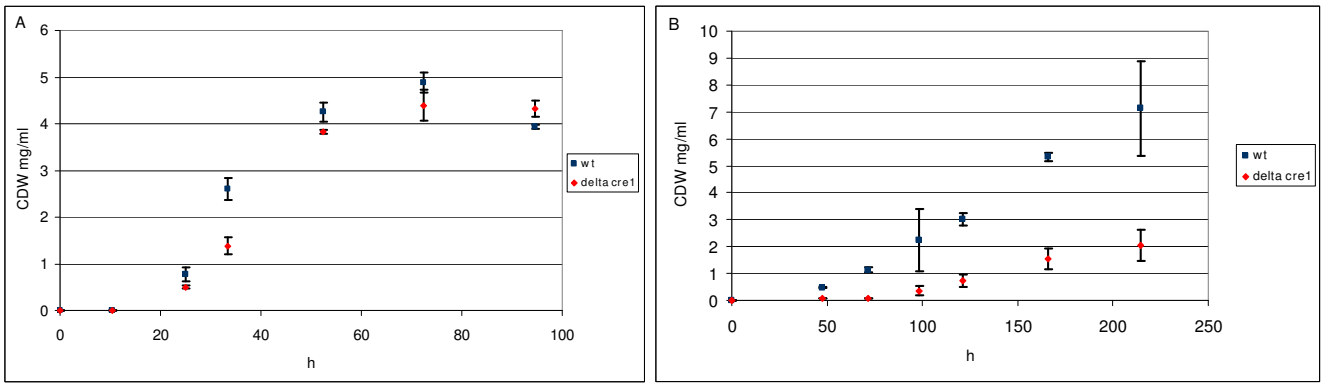

Figure 5
***T. reesei* growth curves**. Growth curves of *T. reesei* wild type (wt) and *Δcre1* strains (A) on glucose and (B) on sorbitol. Error bars are standard deviations of three replicates.

the information of the pathways that were active when the cellulase gene expression was induced was recorded in the labelling patterns of proteinogenic amino acids.

Nuclear Magnetic Resonance (NMR) spectroscopy experiments

The filtered mycelial samples were suspended into 10 ml of 6 M HCl and the biomass was hydrolysed in sealed glass tubes at +110°C for 22 h. The suspensions were dried and dissolved in H_2O for filtration through 0.2 μm filters. The filtrates were vacuum-dried and dissolved in D_2O for NMR experiments. The pH of the samples was below 1 due to residual HCl.

^{13}C-HSQC NMR spectra were acquired at +40°C on a Varian Inova spectrometer operating at a ^{1}H-resonance frequency of 600 MHz essentially as described [22]. For each sample two spectra were acquired focusing on the aliphatic and aromatic regions. For the aliphatic spectra, a matrix of 1024 × 1500 (f2 × f1) complex data points was acquired and zero-filled to 4096 complex data points in f1. The spectral widths were 6000 Hz and 5100 Hz in the ^{1}H- and ^{13}C-dimensions, respectively. The narrow spectral width in the ^{13}C-dimension leads back-folding of part of the signals to the empty regions of the spectrum. For the aromatic region, a matrix of 1024 × 800 complex data points was acquired and zero-filled to 2048 complex data points in f1. The spectral widths for the aromatic spectra were 6000 Hz and 2815 Hz in the ^{1}H- and ^{13}C-dimensions, respectively. All spectra were weighted with a cosine function in both dimensions prior to the Fourier transformation. The spectra were processed using the standard Varian spectrometer software VNMR (version 6.1, C).

Metabolic Flux Ratio (METAFoR) analysis

The software FCAL (R.W. Glaser; FCAL 2.3.1) [25] was used for the integration of ^{13}C-scalar fine structures of proteinogenic amino acid carbon signals in the ^{13}C-HSQC NMR spectra and the calculation of relative abundances of intact carbon fragments originating from a single source molecule of glucose. The nomenclature used here for the intact carbon fragments, fragmentomers, has been described previously [22]. Briefly, $f^{(1)}$ represents the fraction of molecules in which the observed carbon atom and the neighboring carbons originate from different source molecules of glucose, $f^{(2)}$ the fraction of molecules in which the observed carbon atom and one of the two neighboring atoms originate from the same source molecule of glucose, and $f^{(3)}$ the fraction of molecules in which the observed carbon atom and both neighboring carbons originate from the same source molecule of glucose. If the observed carbon exhibits significantly different ^{13}C-^{13}C scalar coupling constants with the neighboring carbons, $f^{(2)}$ and $f^{(2*)}$ can be distinguished. The fraction of molecules with a conserved bond between the observed carbon atom and the neighboring carbon with the smaller coupling is represented by $f^{(2)}$. $f^{(2*)}$ then denotes the fraction of molecules where the bond is conserved between the observed carbon and the neighboring carbon with the larger coupling. If the observed carbon is located at the end of a carbon chain, $f^{(1)}$ and $f^{(2)}$ fragmentomers can be observed indicating the conservation of the two terminal carbon fragment of the molecule.

The degree of ^{13}C-labeling of the biomass amino acids was determined from the ^{13}C-scalar fine structures of Leu-Cβ and Val-Cγ_2. The biomass was assumed to be fully produced from the fractionally labelled carbon source in the glucose experiments and in sorbitol experiments without sophorose induction because the dry weight of the inoculum was negligible. For the sorbitol experiments with sophorose induction the fraction of labeled biomass was estimated from the dry weight measurement (data not shown). The model of the central carbon metabolism network used in the METAFoR analysis was the one previously developed for eukaryotic metabolism of the yeast *S. cerevisiae* [26] (Figure 6). Fragmentomer information obtained from proteinogenic amino acids can be traced back to the metabolic intermediates in central carbon metabolism through the amino acid synthesis pathways to assess ratios of intracellular fluxes which merge at a metabolic network junction [26]. The biosynthetic pathways of amino acids in *T. reesei* were reconstructed with carbon path analysis method ReTrace [21] described in the next section.

Carbon path analysis

The reconstruction of amino acid biosynthetic pathways from their precursors in *T. reesei* was performed with ReTrace. ReTrace is a recent computational pathway analysis method [21], which can be queried to discover branching metabolic pathways in a universal metabolic database. Specifically, ReTrace aims to find pathways which transfer as many atoms from source to target metabolites as possible.

The reaction database used in ReTrace analysis was KEGG LIGAND, downloaded in March 2009 [40]. Reaction database contained 7827 reactions and 15400 compounds. *Atom mappings*, that describe how atoms are transferred in a reaction from substrate to product metabolites, were defined for 33795 substrate-product pairs in the RPAIR database, which is a subdatabase of KEGG. All reactions were considered bidirectional. To compute reaction scores, a database consisting of 101136 sequences annotated with an EC number in UniProt version 9.3 [41] was queried with the 9129 protein sequences in *T. reesei* genome [20] by blastp [42] using e-value cutoff 10 to detect remote homologs. Each reaction in the KEGG database was assigned a score by taking the maximum BLAST

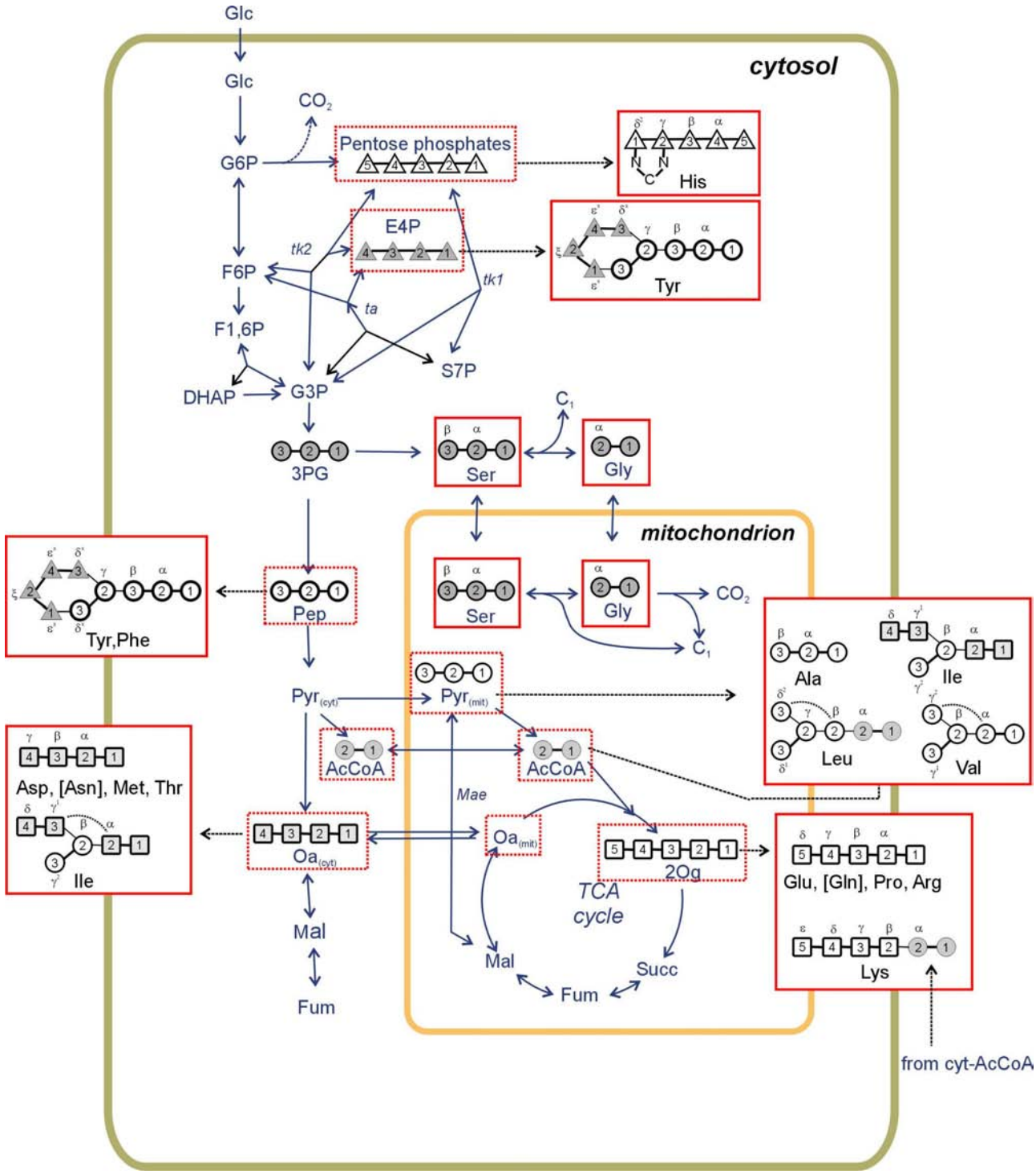

Figure 6
Metabolic network model. Eukaryotic central carbon metabolism network model [26].

score over all UniProt-*Trichoderma* sequence pairs, where UniProt sequence had been annotated with an EC number corresponding to the reaction. A total of 3974 reactions received a score in this procedure, while the remaining 3853 reactions were assigned a zero score. Reaction scores reflected the degree of evidence from the detection of sequence homology that there exists an enzyme catalyzing the reaction in *T. reesei*.

For a majority of pathway queries, maximum search depth was set to 3 and number of pathways searched at depths 1, 2 and 3 to 50, 10 and 1, respectively. In other words, ReTrace search comprised pathways with three branches or less. In particular, more alternative routes (k = 50) were searched at the initial first level (depth 1) than at subsequent levels to reduce the computational complexity. However, in queries involving Asp, Phe, Thr and Tyr, search time with these parameters exceeded a few hours due to branching. These queries were resolved by setting k = 1 already at the second level, while keeping k = 50 at the first level.

Typically, the queries took from 30 minutes to 2 hours CPU time each on computers running Intel Xeon X5355 CPUs. Queries were performed on a cluster of 10 CPUs with four cores each, reducing the total time required. Parameter choices affect the computation time significantly: although it is possible to obtain results on, for example, existence of complete pathways in a matter of seconds by setting k = 1 at each level, in this study a more exhaustive approach was adopted.

Localization of amino acid biosynthetic enzymes in T. reesei

TargetP, a machine learning method based on neural networks, which predicts both chloroplast and mitochondrial targeting peptides and secretory signal peptides, was utilized to predict the probable subcellular localization of some amino acid biosynthetic enzymes in *T. reesei* [33,34]. The prediction performance of non-plant mitochondrial targeting peptides with TargetP has been measured to be 80 - 90% sensitivity and 70% specificity [33]. TargetP reported, for each analyzed peptide sequence, the probability that the peptide contained some signal peptide (SP), a mitochondrial targeting peptide (mTP) or cytosolic targeting peptide (cTP) presequence. In addition, a numerical reliability class (RC) between 1 and 5 was reported. The reliability class was derived from the difference of highest and second-highest probabilities assigned to the prediction classes SP, mTP, cTP or "other". The class "other" indicates the probability that no subcellular location sorting signal was found. If the difference was greater than 0.8, RC equals 1; if the difference was below 0.2, RC equals 5.

Abbreviations

AcCoA: acetyl coenzyme A; Ala L: Alanine; Arg L: Arginine, Asp L: Aspartic acid; BDF: biosynthetically directed fractional; CCR: carbon catabolite repression: E4P: erythrose 4-phosphate; Glu L: Glutamic acid; Gly: Glycine; His L: Histidine; HSQC: heteronuclear single quantum correlation; Ile L: Isoleucine, Leu L: Leucine; Lys L: Lysine; Oaa: oxaloacetate; Oaa_{cyt}: cytosolic oxaloacetate; Oaa_{mit}: mitochondrial oxaloacetate; Oga: oxoglutarate; Pep: phosphoenol pyruvate: Phe L: Phenylalanine; PPP: pentose phosphate pathway; Pro L: Proline; Pyr: pyruvate; Pyr_{cyt}: cytosolic pyruvate: Pyr_{mit}: mitochondrial pyruvate; SEM: standard error of the mean: Ser L: Serine; Thr L: Threonine; Tyr L: Tyrosine.

Authors' contributions

PJ participated in the design of the study, performed the cultivations, carried out the NMR experiments and performed the ^{13}C-metabolic flux ratio analysis, EP performed the ReTrace carbon path analysis, PJ and EP interpreted the results of the computational pathway analysis and wrote the manuscript, HM participated in the design of the study and the NMR experiments, TP, MS and MP participated in the design of the study. All authors read and approved the final manuscript.

Additional material

Additional file 1

Pathways discovered in ReTrace carbon path analysis. *Graphical and tabular representations of amino acid synthesis pathways discovered in ReTrace carbon path analysis [21]. Self-contained web site: unpack zip archive and open index.html with a web browser.*

Click here for file

[http://www.biomedcentral.com/content/supplementary/1752-0509-3-104-S1.zip]

Additional file 2

Estimation of subcellular localization of key enzymes by TargetP. *TargetP machine learning program [33,34] was used to predict the subcellular localization of key enzymes of amino acid biosynthetic routes in* T. reesei. *Targeting sequences were predicted for* T. reesei *sequences that showed the highest homology to the aspartate aminotransferase, acetolactate synthase, serine hydroxymethyltransferase and glycine dehydrogenase in* S. cerevisiae.

Click here for file

[http://www.biomedcentral.com/content/supplementary/1752-0509-3-104-S2.doc]

Acknowledgements

This work was supported by the Academy of Finland (Centre of Excellence, Industrial Biotechnology 2000-2005; projects number 64330, 214568 and 204369, and SYSBIO programme; project number 207435, Centre of Excellence, Algodan 2008-2013; project number 118653 and Centre of Excellence, White Biotechnology - Green Chemistry 2008-2013; project number

118573) and Tekes, the Finnish Funding Agency for Technology and Innovation (Project numbers 40038/4, 40135/04 and 40537/05).

We thank Aili Grundström for expert technical assistance and Marilyn Wiebe and Vytas Raulinaitis for valuable comments.

References

1. Kumar R, Singh S, Singh OV: **Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives.** *J Ind Microbiol Biotechnol* 2008, **35:**377-391.
2. Ilmén M: **Molecular mechanisms of glucose repression in the filamentous fungus *Trichoderma reesei*.** In *PhD thesis* VTT Publications VTT Espoo; 1997.
3. Aro N, Pakula T, Penttilä M: **Transcriptional regulation of plant cell wall degradation by filamentous fungi.** *FEMS Microbiol Rev* 2005, **29:**719-735.
4. Stricker AR, Mach RL, de Graaf LH: **Regulation of transcription of cellulases- and hemicellulases-encoding genes in *Aspergillus niger* and *Hypocrea jecorina* (*Trichoderma reesei*).** *Appl Microbiol Biotechnol* 2008, **78:**211-220.
5. Margolles-Clark M, Ilmén M, Penttilä M: **Expression patterns of ten hemicellulase genes from filamentous fungus *Trichoderma reesei* on various carbon sources.** *J Biotechnol* 1997, **57:**167-179.
6. Ilmén M, Saloheimo A, Onnela M-L, Penttilä ME: **Regulation of Cellulase Gene Expression in the Filamentous Fungus *Trichoderma reesei*.** *Appl Env Microbiol* 1997, **63:**1298-1306.
7. Druzhinina IS, Schmoll M, Seiboth B, Kubicek P: **Global Carbon Utilization Profiles of Wild-Type, Mutant, and Transformant Strains of *Hypcrea jecorina*.** *Appl Env Microbiol* 2006, **72:**2126-2133.
8. Richard P, Putkonen M, Väänänen R, Londesborough J, Penttilä M: **The missing link in the fungal L-arabinose catabolic pathway: identification of the L-xylulose reductase gene.** *Biochemistry* 2002, **41:**6432-6437.
9. Nogawa M, Goto M, Okada H, Morikawa Y: **L-sorbose induces cellulase gene transcription in the cellulolytic fungus *Trichoderma reesei*.** *Curr Genet* 2001, **38:**329-334.
10. Pail M, Peterbauer T, Seiboth B, Hametner C, Druzhinina I, Kubicek CP: **The metabolic role and evolution of L-arabinitol 4-dehydrogenase of *Hypocrea*.** *jecorina* 2004, **271:**1864-1872.
11. Bailey C, Arst HN: **Carbon catabolite repression in *Aspergillus nidulans*.** *Eur J Biochem* 1975, **51:**573-577.
12. Strauss J, Mach RL, Zeilinger S, Hartler G, Stöffler G, Wolschek M, Kubicek CP: **Cre1, the carbon catabolite repressor protein from *Trichoderma reesei*.** *FEBS Lett* 1995, **376:**103-107.
13. Ilmén M, Thrane C, Penttilä M: **The glucose repressor gene *cre1* of *Trichoderma*: Isolation and expression of a full length and truncated mutant form.** *Mol Gen Genet* 1996, **251:**451-460.
14. Cziferszky A, Mach RL, Kubicek CP: **Phosphorylation Positively Regulates DNA Binding of the Carbon Catabolite Repressor Cre1 of *Hypocrea jecorina (Trichoderma reesei)*.** *J Biol Chem* 2002, **277:**14688-14694.
15. Vautard G, Cotton P, Fèvre M: **The glucose repressor CRE1 from *Sclerotina sclerotiorum* is functionally related to CREA from *Aspergillus nidulans* but not to the Mig proteins from *Saccharomyces cerevisiae*.** *FEBS Letters* 1999, **453:**54-58.
16. Pfeiffer T, Schuster S, Bonhoeffer S: **Cooperation and Competition in the Evolution of ATP-producing Pathways.** *Science* 2001, **292:**504-507.
17. Chambergo FS, Bonaccorsi ED, Ferreira AJS, Ramos ASP, Ribamar Ferreira Júnior J, Abrahao-Neto J, Farah JPS, El-Dorry H: **Elucidation of the Metabolic Fate of Glucose in the Filamentous Fungus *Trichoderma reesei* Using Expressed Sequence Tag (EST) Analysis and cDNA Microarrays.** *J Biol Chem* 2002, **277:**13983-13988.
18. Gancedo JM: **Yeast carbon catabolite repression.** *Microbiol Mol Biol Rev* 1998, **62:**334-361.
19. David H, Krogh AM, Roca C, Åkesson M, Nielsen J: **CreA influences the metabolic fluxes of *Aspergillus nidulans* during growth on glucose and xylose.** *Microbiology* 2005, **151:**2209-2221.
20. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EGJ, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parviainen A, Schoch CL, Yao J, Barbote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS: **Genome sequencing and analysis of the biomass-degrading *Trichoderma reesei* (syn. *Hypocrea jecorina*).** *Nat Biotechnol* 2008, **26:**553-560.
21. Pitkänen E, Jouhten P, Rousu J: **Inferring branching pathways in genome-scale metabolic networks with ReTrace.** *BMC Syst Biol* 2009, **3:**103.
22. Szyperski T: **Biosynthetically directed fractional ^{13}C-labelling of proteinogenic amino acids. An efficient tool to investigate intermediary metabolism.** *Eur J Biochem* 1995, **232:**433-448.
23. Sauer U, Hatzimanikatis V, Bailey JE, Hochuli M, Szyperski T, Wüthrich K: **Metabolic fluxes in riboflavin-producing *Bacillus subtilis*.** *Nat Biotechnol* 1997, **15:**448-452.
24. Sauer U, Lasko DR, Fiaux J, Hochuli M, Glaser R, Szyperski T, Wüthrich K, Bailey JE: **Metabolic flux ratio analysis of genetic and environmental modulations of *Escherichia coli* central carbon metabolism.** *J Bacteriol* 1999, **181:**6679-6688.
25. Szyperski T, Glaser RW, Hochuli M, Fiaux J, Sauer U, Bailey JE, Wüthrich K: **Bioreaction network topology and metabolic flux ratio analysis by biosynthetic fractional ^{13}C labelling and two-dimensional NMR spectroscopy.** *Metab Eng* 1999, **1:**189-197.
26. Maaheimo H, Fiaux J, Ìakar PZ, Bailey JE, Sauer U, Szyperski T: **Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional ^{13}C labelling of common amino acids.** *Eur J Biochem* 2001, **268:**2464-2479.
27. Fiaux J, Ìakar PZ, Sonderegger M, Wüthrich K, Szyperski T, Sauer U: **Metabolic-Flux Profiling of the Yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*.** *Eukaryot Cell* 2003, **2:**170-180.
28. Sola A, Maaheimo H, Ylonen K, Ferrer P, Szyperski T: **Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*.** *Eur J Biochem* 2004, **271:**2462-2470.
29. Xu H, Qian BAJ, West AH, Cook PF: **The α-Aminoadipate Pathway for Lysine Biosynthesis in Fungi.** *Cell Biochem Biophys* 2006, **46:**43-46.
30. Broquist HP: **Lysine biosynthesis (Yeast).** *Methods Enzymol* 1971, **17b:**112-119.
31. Blank LM, Lehmbeck F, Sauer U: **Metabolic-flux analysis of fourteen hemiascomycetous yeasts.** *FEMS Yeast Res* 2005, **5:**545-558.
32. Monschau N, Stahmann K-P, Sahm H, McNeil JB, Bognar AL: **Identification of *Saccharomyces cerevisiae* GLY1 as a threonine aldolase: a key enzyme in glycine biosynthesis.** *FEMS Microbiology Letters* 1997, **150:**55-60.
33. Emanuelsson O, Brunak S, von Heijnen G, Nielsen H: **Locating proteins in the cell using TargetP, SignalP and related tools.** *Nature Protocols* 2007, **2:**953-971.
34. Emanuelsson O, Nielsen H, Brunak B, von Heijnen G: **Predicting subcellular localization of proteins based on their N-terminal amino acid sequenc.** *J Mol Biol* 2000, **300:**1005-1016.
35. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'shea EK: **Global analysis of protein localization in budding yeast.** *Nature* 2003, **425:**686-691.
36. Christensen KE, MacKenzie RE: **Mitochondrial one-carbon metabolism is adapted to the specific needs of yeast, plants and mammals.** *Bio Essays* 2006, **28:**595-605.
37. McNeil JB, McIntosh EM, Taylor BV, Zhang F, Tang S, Bognar AL: **Cloning and Molecular Characterization of Three Genes, Including Two Genes Encoding Serine Hydroxymethyltransferases, Whose Inactivation Is Required to Render Yeast Auxotrophic for Glycine.** *J Biol Chem* 1994, **269:**9155-9165.
38. McNeil JB, Bognar AL, Pearlman RE: ***In vivo* analysis of folate coenzymes and their compartmentation in *Saccharomyces cerevisiae*.** *Genetics* 1996, **142:**371-381.
39. Mandels M, Reese ET: **Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals.** *J Bacteriol* 1957, **73:**269-278.
40. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y: **KEGG for linking genomes to life and the environment.** *Nucleic Acids Res* 2008, **36:**D480-D484.

41. The UniProt Consortium: **The Universal Protein Resource (UniProt).** *Nucleic Acids Res* 2007, **35:**D193-197.
42. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25:**3389-3402.
43. Eppstein D: **Finding the k shortest paths.** *35th IEEE Symp Foundations of Comp Sci., Santa Fe* 1994:154-165.

PUBLICATION IV

An analytic and systematic framework for estimating metabolic flux ratios from ^{13}C tracer experiments

In: BMC Bioinformatics 2008. 9:266.

BMC Bioinformatics

BioMed Central

Methodology article

Open Access

An analytic and systematic framework for estimating metabolic flux ratios from ^{13}C tracer experiments

Ari Rantanen*[1,2], Juho Rousu[1], Paula Jouhten[3], Nicola Zamboni[2], Hannu Maaheimo[3] and Esko Ukkonen[1]

Address: [1]Department of Computer Science, University of Helsinki, Finland, [2]Institute of Molecular Systems Biology, ETH Zurich, Switzerland and [3]VTT Technical Research Centre of Finland, Finland

Email: Ari Rantanen* - ari.rantanen@iki.fi; Juho Rousu - juho.rousu@cs.helsinki.fi; Paula Jouhten - pjouhten@mappi.helsinki.fi; Nicola Zamboni - zamboni@imsb.biol.ethz.ch; Hannu Maaheimo - hannu.maaheimo@helsinki.fi; Esko Ukkonen - ukkonen@cs.helsinki.fi

* Corresponding author

Published: 6 June 2008

BMC Bioinformatics 2008, **9**:266 doi:10.1186/1471-2105-9-266

Received: 14 January 2008
Accepted: 6 June 2008

This article is available from: http://www.biomedcentral.com/1471-2105/9/266

Abstract

Background: Metabolic fluxes provide invaluable insight on the integrated response of a cell to environmental stimuli or genetic modifications. Current computational methods for estimating the metabolic fluxes from ^{13}C isotopomer measurement data rely either on manual derivation of analytic equations constraining the fluxes or on the numerical solution of a highly nonlinear system of isotopomer balance equations. In the first approach, analytic equations have to be tediously derived for each organism, substrate or labelling pattern, while in the second approach, the global nature of an optimum solution is difficult to prove and comprehensive measurements of external fluxes to augment the ^{13}C isotopomer data are typically needed.

Results: We present a novel analytic framework for estimating metabolic flux ratios in the cell from ^{13}C isotopomer measurement data. In the presented framework, equation systems constraining the fluxes are derived automatically from the model of the metabolism of an organism. The framework is designed to be applicable with all metabolic network topologies, ^{13}C isotopomer measurement techniques, substrates and substrate labelling patterns.

By analyzing nuclear magnetic resonance (NMR) and mass spectrometry (MS) measurement data obtained from the experiments on glucose with the model micro-organisms *Bacillus subtilis* and *Saccharomyces cerevisiae* we show that our framework is able to automatically produce the flux ratios discovered so far by the domain experts with tedious manual analysis. Furthermore, we show by *in silico* calculability analysis that our framework can rapidly produce flux ratio equations – as well as predict when the flux ratios are unobtainable by linear means – also for substrates not related to glucose.

Conclusion: The core of ^{13}C metabolic flux analysis framework introduced in this article constitutes of flow and independence analysis of metabolic fragments and techniques for manipulating isotopomer measurements with vector space techniques. These methods facilitate efficient, analytic computation of the ratios between the fluxes of pathways that converge to a common junction metabolite. The framework can been seen as a generalization and formalization of existing tradition for computing metabolic flux ratios where equations constraining flux ratios are manually derived, usually without explicitly showing the formal proofs of the validity of the equations.

Background

From microorganisms to animals and plants, cells adjust their metabolic operations to fulfill the demand of energy and biosynthetic precursors. In nature this is a challenging task because substrate availability is typically limited and often changing in its composition. To ensure viability on a broad palette of chemically heterogeneous substrates, cells have developed intertwined enzymatic networks that also confer robustness against genetic mutations. Optimum redistribution of molecular fluxes in metabolism is achieved by multilevel regulation circuits. In recent years, experimental measurement of in vivo metabolic fluxes has attracted much attention. For example, in biotechnology metabolic fluxes are utilized to lead rational strain engineering, whereas systems biologists assess fluxes to unravel targets and mechanisms of metabolic regulation.

Metabolic fluxes are often estimated using flux balance analysis (FBA) [1,2]. In FBA, fluxes are solved by fixing some objective for the metabolism of an organism, such as maximal growth. Then, a corresponding optimization problem is solved by using the stoichiometry of the metabolic network as a constraint to the optimization. FBA is a viable method for studying the metabolic capabilities of an organism, but as a method for estimating metabolic fluxes it has some weaknesses. First, selecting the correct objective for the metabolism is far from trivial [3], especially when mutant strains or behaviour in exceptional environmental conditions is analyzed. Second, there can be many biologically interesting flux distributions that give an optimal solution to the optimization problem of FBA.

A more direct method for experimental determination of the metabolic fluxes is to feed an organism with ^{13}C labelled substrate, observe the fate of ^{13}C atoms in the cell at isotopomeric steady state with mass spectrometry (MS) or nuclear magnetic resonance (NMR) measurements, and then infer the metabolic fluxes from the measurements. The rationale behind these ^{13}C tracer experiment is that, often alternative pathways between metabolites in the network manipulate the carbon backbones of the metabolites differently, thus inducing different ^{13}C labelling patterns to metabolites. Then, constraints to fluxes complementary to the basic stoichiometric constraints can be derived by measuring the relative abundances of different labelling patterns in the metabolites.

The main difficulty in applying the procedure in practice is that current measurement techniques only can produce incomplete information about relative abundances of different ^{13}C labelling patterns, the *isotopomer distributions*, of some metabolites, usually protein bound amino acids in the network, and no isotopomer information at all for many intermediate metabolites of interest [4-6]. This imposes a highly non-linear dependency between the measured isotopomer distributions of the metabolites and the metabolic fluxes, which is very challenging to solve both computationally and statistically.

Currently, two main approaches for ^{13}C metabolic flux analysis exist. In the *global isotopomer balancing* approach, the problem of estimating metabolic fluxes from the isotopomer measurements is formulated as a nonlinear optimization problem, where candidate flux distributions are iteratively generated until they fit well enough to the experimental ^{13}C labelling patterns [7-11]. Global isotopomer balancing is a versatile approach that can be applied with all network topologies, substrate labelling distributions and with all measurement techniques – also in short time scales where isotopomeric steady state is not reached [12-14]. However, due to the nonlinearity of the problem, it is hard to make sure that the optimization has converged to a global optimum and that this optimum is unique [15]. Also, to apply the global isotopomer balancing approach successfully, one usually needs comprehensive information on the uptake and production rates of external metabolites, as well as about biomass composition of the cell. This information can be hard to obtain, especially in large-scale experiments where dozens to hundreds of mutants or different organisms are comparatively analyzed [16,17].

A metabolic flux ratio approach (METAFoR) [4,18,19] for ^{13}C metabolic flux analysis has traditionally relied more on the expertise of a domain specialist than advanced computational techniques. In metabolic flux ratio analysis, the aim is to write linear equations constraining the *ratios* of fluxes producing the same metabolite. The equations are manually derived by domain experts, by careful (and tedious) inspection of metabolic networks. The motivation for the approach is that, in many cases, the knowledge about the flux ratios already offers enough information about the response of an organism to its environment.

The ratio of competing fluxes or pathways producing the same metabolite is easy to understand, and in many cases easier to estimate reliably than all the fluxes in the network – some interesting flux ratios might be obtainable from scarce measurement data or from the incomplete model of metabolic network that would not allow a reliable estimation of a complete flux distribution using global isotopomer balancing. Flux ratios can also be obtained without knowing the uptake and production rates of external metabolites of the biomass composition. And, if enough non-redundant flux ratios are identified, it is possible to use this information to construct and solve an equation system for the full flux distribution of the metabolic network [20-22].

As a downside, manually derived flux ratio equations depend heavily on the topology of a metabolic network, measurement capabilities and substrate labelling distributions. Thus, each time a new organism or new mixture of substrates are introduced, flux ratio equations have to be verified and new ones possibly derived. To date, flux ratio equations are manually derived for central carbon metabolisms of three model organisms on glucose, *S. cerevisiae* [17,23], *B. subtilis* [24] and for *Escherichia coli* [18,19]. Recently, flux ratio equations of *S. cerevisiae* were modified for *Pichia pastoris* grown on glycerol and on glycerol/methanol mixtures [25,26]. Facilitating the process of deriving flux ratio equations for other organisms and other substrates clearly calls for automatic tools. Also, many times the (simplifying) assumptions made by the expert in the derivation and solution of flux ratio equations, are not reported in detail. Thus, it is often nontrivial to verify the correctness of given flux ratio equations.

In this article we present a novel automatic framework for deriving flux ratios when the measurement data and the model of metabolic network are given as input. The framework is based on the graph algorithmic flow analysis of metabolite fragments in the metabolic network [27] and on the interpretation and manipulation of both NMR and MS data with vector space techniques [21]. The goal of our work is to combine the good aspects of global isotopomer balancing and manual flux ratio analysis: like global isotopomer balancing, our framework is systematic and can be applied with all network topologies, substrates and substrate labelling distributions and with all current isotopomer measurement techniques. Thus, laborious and error-prone manual inspection of metabolic network models and the tailoring of the equation systems constraining the fluxes separately for each experimental setting required in manual flux ratio techniques can be avoided. On the other hand, during the automated construction of flux ratios we resort to linear optimization techniques only, combined with graph algorithms of polynomial worst case time complexity. Thus, our framework is computationally efficient and avoids problems with local and multiple optima frequently met in global isotopomer balancing. The trade-off of this philosophy is, however, the requirement of measuring isotopomer distributions of metabolites more rigorously to obtain full flux distribution. Given insufficient measurements, our framework can solve the flux ratios only for some, but not necessary for all metabolites in the network. We expect that, especially as measuring isotopomers of intermediate metabolites becomes more routine, our framework will be an attractive method for ^{13}C flux analysis.

Results and Discussion

In this section we demonstrate the applicability of the presented framework by empirical results. We show that our automatic and systematic framework is able to reproduce flux ratios previously determined by a manual analysis from NMR and GC-MS isotopomer measurements of protein bound amino acids of *S. cerevisiae* and *B. subtilis* on glucose. Thus, we can conclude that the presented framework is powerful enough to provide interesting flux ratio information in the well studied experimental settings. Furthermore, we show that the framework can be applied to study less known experimental conditions without any further effort by discovering the flux ratios that can be estimated when *B. subtilis* is grown on malate instead of glucose. The results of this analysis show that our framework can detect profound effects the change of external substrate can have to the flux ratio computations. The results indicate that our framework is a good tool to study flux ratios of microbes in different experimental conditions – a claim that will try to validate with more experiments in our further work.

We obtained NMR and GC-MS labelling data, where isotopomer distributions of protein bound amino acids of *S. cerevisiae* and *B. subtilis* grown on different conditions were measured. Then, available flux ratios were computed with the presented framework. Models of metabolic networks applied in the analysis can be found from the supplementary material of this article: additional files 1 and 2 contain the SBML model file and a visualization of the model of *S. cerevisiae*, while additional files 3 and 4 contain the same information for *B. subtilis*. In the models, some simplifications common to ^{13}C metabolic flux analysis were applied by pooling metabolites whose isotopomer pools can be assumed to be fully mixed (cf. [28]). Pooling of metabolites was carried for (i) the three pentose-phosphates in PPP, (ii) phopshotrioses between GA3P and PEP in glycolysis, and (iii) oxaloacetate and malate in the TCA cycle. In these cases, pooling is justified by the existence of fast equilibrating, bidirectional reactions between the listed intermediates and the empirical evidence that their isotopic labeling is not distinguishable with the current analytical tools. Cofactor metabolites were excluded from the model as cofactor specificities and activities are not accurately known for many reactions.

The bulk of the carbon mappings of reactions in the metabolic network were provided by ARM project [29]. Carbon mappings from amino acids to their precursors were conform to [4] and [23]. Before the analysis of the real measurement data, the correctness of the implementation of the framework was empirically verified by estimating flux ratios for junction metabolites in the metabolic network of S. cerevisiae from the artificial data generated by the 13C-FLUX software [8].

NMR measurements from S. cerevisae *on glucose*

In the first experiment we analyzed NMR isotopomer measurement data from protein bound amino acids of *S. cerevisiae* that was grown on uniformly labelled glucose (see Section *Experimental NMR and GC-MS methods* for more details on experimental settings).

From the 15 measured amino acids we were able to estimate flux ratios for seven junction metabolites: oxaloacetate, PEP, glycine and serine on cytosol and for oxaloacetate, acetyl-CoA and pyruvate in mitochondria. Furthermore, an upper bound for a ratio of GA3P molecules that have visited the transketolase reaction was obtained by manually simplifying the model to imitate the previously reported ways to manually compute the corresponding upper bound (cf. [4]). (The structural analysis of the metabolic network model described in Section *Structural analysis of isotopomer systems* can help in discovering such simplifications, but they also need some expert insight. As the simplifications are currently not done automatically, the systematical framework is unable to validate them.)

The computed flux ratios were compared with the manually derived ratios [23], when the assumptions made in the manual derivation of flux ratios were consistent with the model used. In all cases, automatically computed flux ratios agreed well with the manually derived ratios (Table 1). Differences between the estimations can be explained by numerical instabilities and by differences in computational procedures: in manually derived ratios the estimations are based on the breakage of a single bond in different routes leading to a metabolite while in our framework more isotopomer information is possible utilized in the estimation.

GC-MS measurements from B. subtilis *on glucose*

In the second experiment we analyzed GC-MS isotopomer measurement data from protein bound amino acids of *Bacillus subtilis* that was grown on uniformly labelled glucose (see Section *Experimental NMR and GC-MS methods* for more details on experimental settings).

In comparison to eukaryotic *S. cerevisiae*, the metabolic network of prokaryotic *B. subtilis* lacks cellular compartments. Thus, from the point of view of ^{13}C metabolic flux analysis, there are fewer interesting junction metabolites in the central carbon metabolism of *B. subtilis* where the flux ratios can be estimated. From the GC-MS measurements of 14 amino acids we were able to compute flux ratios for four junction metabolites – oxaloacetate, pyruvate, PEP and glycine – when [U-13C]-glucose was used as a carbon source. Furthermore, an upper bound for a ratio of GA3P molecules that have visited transketolase reaction was obtained by manually simplifying the model of the metabolic network. Excluding pyruvate, we were able to compute the same ratios with [1-13C]-glucose as a carbon source.

We compared the computed flux ratios to ones obtained with the software FiatFlux [30] that is based on the manually derived analytic equations for computing flux ratios. Currently, manually derived flux ratio equations for [1-13C]-glucose as a carbon source exist only for PEP and for the upper bound to the flux through oxidative pentose phosphate pathway. In general, the flux ratios computed with different methods from the same data and with the same assumptions about the topology of metabolic network were in good agreement (Table 2). (As a data cleaning procedure, we removed from [1-13C]-glucose data the mass distributions of fragments whose fractional enrichment deviated more than 5% from the assumed fractional enrichment of 20% in [U-13C]-data. This was done because differences in fractional enrichments can be tracked in uniformly labelled data where the fractional enrichment of each fragment is know *a priori*, but not in positionally labelled data, where the fractional enrichment of a fragment depends on the network topology and the fluxes. This kind of irregularities are in general caused by noise in fragments with low intensity or by coeluting analytes with overlapping fragment masses.) Again, differences between the flux ratios estimated by different methods can be explained by numerical instabilities and by differences in isotopomer information applied during the estimation. Variation in the estimated flux ratios between repeated experiments (six repetitions for [1-13C]-glucose

Table 1: Estimated flux ratios from NMR measurements of *S. cerevisiae*.

flux ratios	our framework	METAFoR
PYR(mit) from MAL(mit) : PYR(mit) from PYR(cyt)	0.05 : 0.95	0.03 : 0.97
OAA(mit) from TCA-cycle : OAA(mit) from OAA(cyt)	0.50 : 0.50	0.50 : 0.50
PEP from OAA(cyt) : PEP from GA3P	0 : 1	0.04 : 0.96
OAA(cyt) from PYR(cyt) : OAA(cyt) from OAA(mit)	0.40 : 0.60	0.43 : 0.57
GLY from SER : GLY from C1 + CO2	0.96 : 0.04	0.96 : 0.04

Directly comparable flux ratios computed from the NMR data by the framework presented in this paper and by manual flux ratio analysis (METAFoR).

Table 2: Estimated flux ratios from GC-MS measurements of *B. subtilis*.

flux ratios	our fw (UL)	our fw (1CL)	FiatFlux (UL)	FiatFlux (1CL)
PYR from MAL : PYR from PEP	0.01 : 0.99		0.04 : 0.96	
OAA from TCA-cycle : OAA from PYR	0.42 : 0.58	0.37 : 0.63	0.41 : 0.59	
PEP from OAA : PEP from GA3P	0 : 1	0 : 1	0.04 :0.96	0 : 1
GLY from SER : GLY from C1 + CO2	1 : 0	1 : 0	1 : 0	
SER from GLY : SER from GA3P	0.09 : 0.91		0.14 : 0.86	

Directly comparable flux ratios computed from the GC-MS data described by the framework presented in this paper (*our fw*) and by FiatFlux software [30]. *UL* denotes uniform labelling of external glucose, *1CL* external glucose labelled to the first carbon position.

experiment, four repetitions for uniformly labelled glucose experiment) was negligible.

In silico calculability analysis of B. subtilis *on malate*

One of the strengths of the presented framework is that it is able to automatically produce metabolic flux ratios also when other external labelled substrates than commonly used glucose are fed to organisms. To demonstrate this ability, we applied our framework to predict what flux information would be available, if we feed *B. subtilis* with malate.

We applied the *in silico* calculability analysis (see Section *Calculability analysis*) to examine which flux ratios are calculable in the best case from GC-MS measurements of amino acids, when *B. subtilis* is grown on [U-13C]-labelled malate. Interestingly, our fragment flow analysis revealed that – with the reaction reversibilites in the applied model – the isotopomer distributions of GA3P, PEP and pyruvate depend only on the isotopomer distribution of the fragment containing the first three carbons of oxaloacetate, but not on the relative fluxes producing these metabolites. Thus, isotopomer balances for GA3P, PEP and pyruvate reduce to mass balances and the ratios of fluxes producing these metabolites cannot be estimated. This somewhat surprising phenomenon is due the fact that the rearrangements of carbon chains occurring in PPP pathway will affect only to the carbon fragments that will be recycled in the upper metabolism but not the carbon fragments that can flow back GA3P, PEP and pyruvate from PPP (we modelled a reaction from GA3P to F6P as unidirectional one).

Preliminary experiments with GC-MS data from *B. subtilis* grown on [U-13C]-labelled malate agreed with the results of fragment flow analysis: constraints to the isotopomer distributions of fluxes entering to these metabolites were identical within the limits of measurement accuracy. On the other hand, our framework was able to estimate for example the TCA-cycle activity also when *B. subtilis* is grown on malate, just as predicted by the calculability analysis.

Conclusion

In this article we introduce a systematic and analytic framework for ^{13}C metabolic flux ratio analysis. At the heart of the framework lie the techniques for flow analysis of a metabolic network and for manipulating isotopomer measurements as linear subspaces. These techniques facilitate the efficient and analytic computation of the ratios between the fluxes producing the same junction metabolite. The framework can be seen as a generalization and formalization of existing analytic methods for computing metabolic flux ratios [23,30,31] where equations constraining flux ratios are manually derived. Like the recent methods to improve the speed of the simulation of isotopomer distributions in the global isotopomer balancing framework [10,32], our framework relies on graph algorithms. However, both our goals and applied techniques are quite different from these approaches. In [10] and [32] connected components of isotopomer graphs are discovered to divide the simulation of isotopomer distributions to smaller subtasks. In our framework, flow analysis techniques are applied to discovered metabolite fragments with equivalent isotopomer distributions in every isotopomeric steady state.

Our experiments with NMR and MS data show that the framework is able to provide relevant information about metabolic fluxes, even when only constraints to the isotopomer distributions of protein-bound amino acids are measured.

Thanks to recent advancements in measurement technology improving the feasibility of mass isotopomer measurements of intermediate metabolites [13,33], we expect that the full power of the framework will be harnessed in near future. Measurements from intermediates will make it possible to use larger models of metabolic networks and estimate flux ratios more accurately, without simplifying assumptions about the topology of the metabolic network or directionality of the fluxes. However, these measurements will not be easy to conduct, because of the low abundances of intermediates in the cell. Thus, systematic methods for experimental planning and data quality control are required. The presented framework provides pow-

erful tools for these tasks. First, the framework facilitates time saving *in silico* calculability analysis.

Second, the manipulation of isotopomer measurements as linear subspaces offers a natural way for comparing measurements from different metabolites that contain overlapping information to detect inconsistencies in the measurements: it is enough to compare propagated isotopomer information in the fragments that belong to the same equivalence class. Third, as MS isotopomer measurement techniques have to be developed separately for different intermediate metabolites or metabolite classes, it will be very useful to select a small subset of intermediates that gives enough information about the interesting metabolic fluxes with least experimental effort. In future research, we want to tackle this problem by generalizing our earlier experimental planning method [34], to all measurement data and to realistic measurement error models.

As the presented framework for ^{13}C metabolic flux analysis only resorts to linear optimization techniques, it is not always able to provide as much information about the metabolic fluxes as the global isotopomer balancing frameworks utilizing more powerful, nonlinear optimization techniques [8,35], that do not necessarily converge to the global optimum. On the other hand, some flux ratios might be computable from the scarce data or incomplete model of metabolic network that does not allow global isotopomer balancing. The differences in the practical performance of different approaches require further research. We see these alternative approaches as complementary ones. A very nice goal would be an integration of our work with global isotopomer balancing: our analytic flux ratios could speed up and direct the optimization process of global isotopomer balancing, that would then fill in the flux ratios possibly missed by our framework.

Methods

In this section we describe a systematic and analytic computational framework for ^{13}C metabolic flux analysis. At the end of the section we also shortly describe the experimental method that were applied to produce the isotopomer measurement data that was analyzed in Section *Results and Discussion*.

The overall goal of our computational framework is to automatically infer from the available isotopomer measurement data produced by MS, MS-MS or NMR techniques, an equation system constraining the fluxes. The crucial question is to derive as many non-redundant equations as possible, ideally constraining the flux distribution to a point solution, or in general, as low-dimensional convex set as possible.

In short, the framework consists of the following steps:

1. The model of the metabolic network of an organism is constructed by selecting a set of biochemical reactions operating in the organism and by designating them to correct cellular compartments;

2. Structural analysis of the isotopomer system is conducted, consisting of the following steps:

(a) Flow analysis of the metabolic network is conducted in order to discover *equivalent fragments*, fragments of carbon backbones of metabolites that will have the same theoretical isotopomer distribution, regardless of the fluxes.

(b) Independence analysis of fragments is conducted to find statistically independent carbon subsets from metabolites, that is, subsets that have been at some point separated along every pathway able to producing them and that have flux invariant isotopomer distributions. This guarantees that the isotopomer distribution of their union assumes the form of a product distribution.

(c) *In silico* calculability analysis is performed to test if the available measurement techniques and substrate labellings make it *in principle* possible to obtain the required flux information.

3. Wet-lab isotopomer tracer experiments are conducted and constraints to isotopomer distributions are measured;

4. The fluxes of the network are estimated. The process consists of the following steps:

(a) Isotopomer measurement data is *propagated* in the metabolic network model from the measured metabolites to unmeasured ones according to the equivalences discovered in step 2.

(b) An equation system tying the isotopomer data and the fluxes together is constructed and solved, either to obtain a flux distribution for the metabolic network as a whole, or for a single junction metabolite to obtain the ratios of fluxes producing it.

(c) The statistical analysis of obtained fluxes or flux ratios is carried out.

In the following we first formalize the ^{13}C flux analysis problem and then detail the computational steps above.

Preliminaries

In ^{13}C metabolic flux analysis the carbon atoms of metabolites are of special interest. We denote with M the set of carbon locations $M = \{c_1, ..., c_k\}$ of a k-carbon metabolite.

By $|M| = k$ we denote the number of carbons in M. *Fragments* of metabolites are subsets $F = \{f_1, ..., f_h\} \subseteq M$ of the carbons of the metabolite. A fragment F of M is denoted as $M|F$. A *metabolic network* $G = (\mathcal{C}, \mathcal{R})$ is composed of a set $\mathcal{C} = \{M_1, ..., M_m\}$ of *metabolites* and a set $\mathcal{R} = \{\rho_1, ..., \rho_n\}$ of *reactions* that perform the interconversions of metabolites.

With *isotopomers* we mean molecules with similar element structure but different combinations of ^{13}C labels. Isotopomers of the molecule $M = \{c_1, ..., c_k\}$ are represented by binary sequences $b = \{b_1, ..., b_k\} \in \{0, 1\}^k$ where $b_i = 0$ denotes a ^{12}C and $b_i = 1$ denotes a ^{13}C in location c_i. Molecules that belong to the *b-isotopomer* of M are denoted by $M(b)$. Isotopomers of metabolite fragments $M|F$ are defined in an analogous manner: a molecule belongs to the $F(b)$-*isotopomer* of M, denoted $M|F(b_1, ..., b_h)$, if it has a ^{13}C atom in all locations f_j that have $b_j = 1$, and ^{12}C in other locations of F. Isotopomers with equal numbers of labels belong to the same *mass isotopomer*. We denote *mass isotopomers* of M by $M(+p)$, where $p \in \{0, ..., |M|\}$ denotes the number of labels in isotopomers belonging to $M(+p)$.

The *isotopomer distribution* D_M of metabolite M gives the relative abundances $0 \le P_M(b) \le 1$ of each isotopomer $M(b)$ in the pool of M such that

$$\sum_{b \in \{0,1\}^{|M|}} P_M(b) = 1.$$

The isotopomer distribution $D_{M|F}$ of fragment $M|F$ and the *mass isotopomer* distribution D_M^m of mass isotopomers $M(+p)$ are defined analogously: $D_{M|F}$ of metabolite M gives the relative abundances $0 \le P_{M|F}(b) \le 1$ of each isotopomer $M|F(b)$ and $(\rho_1^p, \rho_2^p, ..., \rho_n^p)$ gives the relative abundances $0 \le P_M(+p) \le 1$ of each mass isotopomer $M(+p)$.

Reactions are pairs $\rho_j = (\alpha_j, \lambda_j)$ where $\alpha_j = (\alpha_{1j}, ..., \alpha_{mj}) \in \mathbb{Z}^m$ is a vector of *stoichiometric coefficients*-denoting how many molecules of each kind are consumed and produced in a single reaction event-and λ_j is a carbon mapping describing the transition of carbon atoms in ρ_j (see Figure 1). Metabolites M_i with $\alpha_{ij} < 0$ are called *substrates* and with $\alpha_{ij} > 0$ are called *products* of ρ_j. If a metabolite is a product of at least two reactions, it is called a *junction*. If $\alpha_{ij} < 0$, a reaction event of ρ_j consumes $|\alpha_{ij}|$ molecules of M_i, and if $\alpha_{ij} > 0$, it produces $|\alpha_{ij}|$ molecules of M_i. Bidirectional reactions are modelled as a pair of reactions.

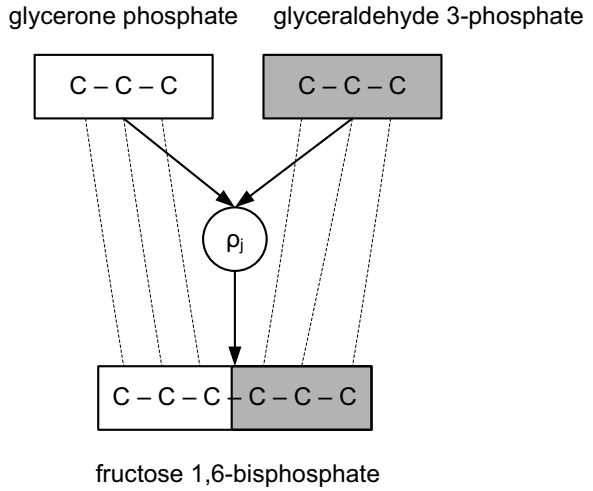

Figure 1
An example of a metabolic reaction. In the reaction ρ_j, a fructose 1,6-bisphosphate ($C_6H_{14}O_{12}P_2$) molecule is produced from glycerone phosphate ($C_3H_7O_6P$) and glyceraldehyde 3-phosphate ($C_3H_7O_6P$) molecules. Carbon maps are shown with dashed lines. Glyceraldehyde 3-phosphate is equivalent to the gray fragment of fructose 1,6-bisphosphate while glycerone phosphate is equivalent to the white fragment.

A *pathway* p in network G from metabolite fragments $\{F_1, ..., F_k\}$ to fragment F' is a sequence $(\rho_1^p, \rho_2^p, ..., \rho_n^p)$ of reactions such that a *composite carbon mapping* $(\lambda_n^p \circ \dots \lambda_2^p \circ \lambda_1^p)(\bigcup_{i \in [1,k]} F_i) = F'$, defined by p maps the carbons of $\{F_1, ..., F_p\}$ to the carbons of F'.

For the rest of the article, it is important to distinguish between the *subpools* of a metabolite pool produced by different reactions. Therefore, we denote by M_{ij}, the subpool of the pool of M_i produced ($\alpha_{ij} > 0$) or consumed ($\alpha_{ij} < 0$) by reaction ρ_j. The concept of the subpools of a metabolite pool is illustrated in Figure 2.

By M_{i0} we denote the subpool of M_i that is related to the external in flow or external out flow of M_i. We call the sources of external in flows *external substrates*. Subpools of fragments are defined analogously. In ^{13}C metabolic flux analysis, the quantities of interest are the rates or the fluxes $v_j \ge 0$ of the reactions ρ_j, giving the number of reaction events of ρ_j per time unit. We denote by $\mathbf{v}$ the vector $[v_1, ..., v_n]$ of fluxes, or a *flux distribution*.

Generalized isotopomer balances

The framework for ^{13}C metabolic flux analysis presented in this article rests on the assumption that the metabolic network is in metabolic and isotopomeric steady state. In

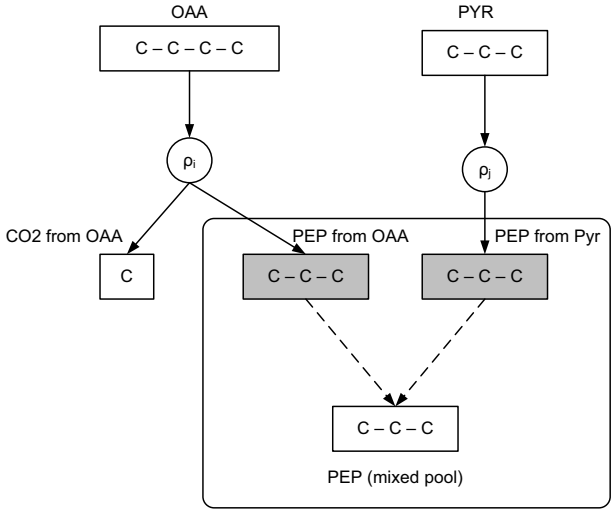

Figure 2
An example of subpools of a metabolite pool. Phosphoenolpyruvate (PEP) is produced by two different reactions (ρ_i and ρ_j), either from Oxaloacetate (OAA) or from glyceraldehyde 3-phosphate (GA3P). Thus, PEP has two in flow subpools, *PEP from OAA* and *PEP from GA3P* (grey boxes) that are mixed in the common PEP pool (at the bottom).

the metabolic steady state, *metabolite balance*, or *mass balance*

$$\sum_{j=1}^{n} \alpha_{ij} v_j = \beta_i \qquad (1)$$

holds for each metabolite M_i. Here, β_i is the measured external in flow ($\beta_i < 0$) or external out flow ($\beta_i > 0$) of metabolite M_i. From balance equations (1) defined for every metabolite M_i we will obtain a metabolite balancing, or stoichiometric equation system.

In isotopomer steady state, for each isotopomer b of each metabolite M_i in the metabolic network the following *isotopomer balance* holds:

$$\sum_{j=1}^{n} \alpha_{ij} v_j P_{M_{ij}}(b) = \beta_i P_{M_{i0}}(b). \qquad (2)$$

For metabolic flux analysis (1) and (2) bear a fundamental difference: the former cannot be used to estimate fluxes of alternative pathways producing M_i while the latter can. However, using (2) is not in general admissible in practice: typically abundances $P_M(b)$ of isotopomers are not fully determined from the measurements, and we need to settle for some constraints to the distribution D_M. A crucial building block of our framework is the representation of isotopomer measurements as systems of linear equations (c.f. [21])

$$\sum_b s_{bh} P_M(b) = d_h, h = 1, ..., r \qquad (3)$$

where s_{bh} is the weight of isotopomer b in the h'th constraint, d_h is a value derived from isotopomer measurements, and r is the total number of constraints. We call (3) *isotopomer constraints*. For a k carbon metabolite, 2^k linearly independent isotopomer constraints – one for each isotopomer – are necessary and sufficient to constrain the isotopomer distribution D_M to a point solution. A set of isotopomer constraints has a natural matrix representation $SD_M = \mathbf{d}$, where $S = (s_{bh})_{b,h}$ is a 2^k by r matrix, where $1 \le r \le 2^k$ is the number of isotopomer constraints (the trivial constraint $\Sigma_b P_M(b) = 1$ by definition always holds).

The use of (3) follows from the simple observation that isotopomer balance (2) implies that each linear combination of isotopomers is balanced. Thus, we can write a new balance equation that constrains the fluxes producing M_i as soon as we know the value of the *same* linear combination of isotopomer abundances for each subpool of M_i. We have

$$\sum_{j=1}^{n} \alpha_{ij} v_j d_{ij} = \beta_i d_{i0}, \qquad (4)$$

where each $d_{ij} = \sum_b s_b P_{M_{ij}}(b)$ is a linear combination of the form (3), with coefficients s_b that do not depend on j, i.e. they are the same for each subpool M_{ij}. We call (4) a *generalized isotopomer balance*.

Representing MS and NMR isotopomer measurements as linear constraints

In the following, we will show by examples how MS and NMR data can be represented as isotopomer constraints. Let us first consider mass isotopomer distributions obtained from MS. Here we omit discussion on practicalities such as corrections for natural abundances of ^{13}C isotopes (c.f., [6,36]) and concentrate on the conceptual level. For example, the +2 mass isotopomer of a three carbon metabolite M satisfies

$$P_M(+2) = P_M(011) + P_M(110) + P_M(101)$$

which conforms to (3) by taking $s_{011,2} = s_{110,2} = s_{101,2} = 1$, and $s_{b,2} = 0$ otherwise. Similarly, the coefficients s_{bk} can be derived for all mass isotopomers $+k$, $k = 0, ..., 3$.

Isotopomer data originating from Tandem MS, or MS-MS, falls into the same representation. Consider, for example,

a fragment $M|F$ of a three-carbon metabolite, containing the first and second carbon of M. The following equation holds for the mass isotopomer $M|F(+1)$:

$$P_{M|F}(+1) = P_{M|F}(01) + P_{M|F}(10).$$

The equation can be written in terms of the precursor M, but the exact form of the equation depends on the mode of tandem MS. If the full scan mode is used, we obtain

$$P_{M|F}(+1) = P_M(010) + P_M(011) + P_M(100) + P_M(101),$$

as all precursor molecules M that have exactly one carbon among the first and second location contribute to the fragment mass isotopomer $M|F(+1)$. On the other hand, in the daughter-ion scanning mode a single mass isotopomer, for example $M(+2)$, is selected as the precursor. Then we obtain

$$P_{M|F}(+1) = P_M(011) + P_M(101),$$

as the precursor must always have two ^{13}C atoms in total. We refer the reader to [6,36] for further details. Also NMR ^{13}C isotopomer measurements, where relative intensities of different combinations of ^{13}C and ^{12}C atoms that are coupled to an observed ^{13}C atom are measured, can be expressed as linear combinations of isotopomer abundances. For example, for a three-carbon metabolite M, the following constraints to D_{M_i} can be inferred for the labeling pattern 010:

$$\frac{P_M(010)}{\sum_{b_1,b_3\in\{0,1\}} P_M(b_1 1 b_3)} = d(010)$$

where $d(010)$ is the measured intensity. Rewriting the above as

$$d(010)\cdot \sum_{b\in\{110,011,111\}} P_M(b) = (1 - d(010))\cdot P_M(010)$$

and denoting $s_b = d(010)$, for $b \in \{110, 011, 111\}$, $s_{010} = d(010) - 1$ and $s_b = 0$ for $b \in \{000, 100, 001, 101\}$, the above can be seen to conform to (3). Similar derivation can be used for other isotopomer signals present in the NMR spectrum to obtain the corresponding isotopomer constraints.

Projection of isotopomer measurements to fragments

In our computational framework, it will be necessary to project the measurement data obtained for a metabolite M to its fragments $M|F$ and vice versa. In this projection, we want to avoid any unnecessary loss of measurement information, that is, we want to obtain as many linearly independent constraints to the isotopomer distribution of F as possible. For example, if we have measured that a two-carbon metabolite M has the isotopomer distribution

$$P_M(00) = 0.2,\ P_M(01) = 0.3,\ P_M(10) = 0.4,\ P_M(11) = 0.1$$

and we need to know the isotopomer distribution of the fragment $M|F$ consisting of the first carbon of M, we should obtain

$$P_F(0) = P_M(0*) = P_M(00) + P_M(01) = 0.5,$$
$$P_F(1) = P_M(1*) = P_M(10) + P_M(11) = \mathbf{0}.5.$$

For the general model of isotopomer measurements (4) the projection of measurement information from a metabolite to its fragments can be done by the techniques of linear algebra introduced in [21]. We recapitulate the techniques in the following.

Recall the general form of isotopomer measurement $SD_M = \mathbf{d}$, where S denotes a matrix with 2^k columns, one column for each isotopomer b of k-carbon metabolite M, and each row h of S corresponds to a measured isotopomer constraint (3). The rows of S span a subspace $\mathcal{S}\subseteq\mathcal{I}_M$ in a 2^k dimensional vector space $\mathcal{I}_M$ spanned by all possible isotopomer distributions D_M.

Also the metabolite fragments are naturally represented as vector subspaces. Let U_F denote a matrix with also a column for each isotopomer $M(b)$ and a row for each isotopomer $F(b')$ of $M|F$, that is,

$$U_F(b', b) = \begin{cases} 1 & \text{if } b_j = b'_j \quad \text{for all carbon positions } j \in F_k \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

The rows of U_F span another subspace $\mathcal{U}_F\subseteq\mathcal{I}_M$. Any isotopomer distribution $D_{M|F}$ lies in this subspace, and hence also any isotopomer constraint $S_F D_{M|F} = \mathbf{d}$ for fragment $M|F$ necessarily lies in the same subspace.

In conclusion, the available information about D_M is given as its linear projection onto $\mathcal{S}$, and anything we can express about $D_{M|F}$ in terms of isotopomer constraints is contained within $\mathcal{U}_F$. Thus, any isotopomer constraint for $D_{M|F}$ that we can derive from the measurements can be expressed in terms of the *vector space intersection* $\mathcal{S}\cap\mathcal{U}_F$.

Thus, to obtain isotopomer constraints for fragment $M|F$ from a measurement $SD_M = \mathbf{d}$, we need to compute the

vector space intersection $\mathcal{Y}_{M|F}=\mathcal{S}\cap\mathcal{U}_F$ and project the measurement to $\mathcal{Y}_{i,F}$. This can be done by standard linear algebra (c.f. [21]). This process gives us as output isotopomer constraints of the required form

$$Y_F D_{M|F} = \mathbf{d}_F.$$

Finally, transforming a fragment constraint $Y_F D_{M|F} = \mathbf{d}_F$ into an isotopomer constraint $SD_M = \mathbf{d}$ is easy: we post-multiply the fragment constraint with the matrix $U_{M|F}$: $S = Y_F U_{M|F}$ and $\mathbf{d} = \mathbf{d}_F U_{M|F}$.

Structural analysis of isotopomer systems

The incomplete nature of ^{13}C measurement data requires us to find ways to use the available data the best way possible. The central concept is to find invariants of isotopomer distributions that remain through the pathways, and allow us to trade or *propagate* measurement information from one metabolite to another. This allows us to write or augment generalized isotopomer balances for metabolites for which the isotopomer distributions are not completely determined by measurements. Thus, the fluxes are potentially more tightly pinpointed as well.

In particular, we use two techniques: First, *flow analysis* is used to uncover sets of metabolite fragments that have the same isotopomer distribution regardless of the fluxes. Second, *independence analysis* of fragments is used to uncover situations where two fragments of the same metabolite induce the product distribution for the isotopomer distribution of their union.

Flow analysis of metabolic networks

The goal of the flow analysis [27] is to partition the fragments of the metabolites in the network to *equivalence classes* such that fragments in the same equivalence class have identical isotopomer distributions in every steady state. This can be guaranteed if a fragment is produced from another a such a manner that the carbons within the fragment never depart from each other regardless of the pathway that is being used.

Formally, we say that fragment F' *dominates* fragment F if the following conditions are met

1. F and F' have the equal number of carbons;

2. all carbons of F originate always from the carbons of F';

3. carbon of F' stay connected to each other via all pathways from F' to F;

4. composite carbon mappings are the same in all pathways from F' to F.

Intuitively, a *dominated fragment* (F) is always produced from its *dominator* (F') without manipulating the carbon chain of the fragment. Thus, isotopomer distribution of the dominated fragment does not contain any information about the metabolic fluxes. For a fragment F that has no dominators, the transitive closure of the domination relation corresponds to the class of equivalent fragments in the network.

The simplest example of fragment equivalence is the one between a substrate M_k and product M_i in a single reaction ρ. If the atoms in $M_i|F$ originate from $M_k|F'$, then the fragments $M_k|F'$ in the subpool M_{ij} produced by reaction ρ_j, are equivalent with the fragment $M_i|F$ (Figure 1). Furthermore, if metabolite M_i has only one producing reaction ρ_j, isotopomer distributions of subpool M_{ij} and M_i coincide. Thus, if fragment $M_i|F$ is produced from a single fragment $M_k|F'$ of some substrate M_k of ρ_j, F and $F'$are equivalent. By transitivity, all fragments in the linear pathway are equivalent.

More complicated case of fragment equivalence is found when a fragment of a junction metabolite is dominated by an upstream fragment (Figure 3). In [27] we show that the the classes of equivalent fragments corresponding the conditions (1–4) can be efficiently computed. Very brie fly, first the metabolic network is transformed to a *fragment flow graph* that connects substrate metabolite fragments to their product fragments for each reaction in the network. Then, the dominator tree [37,38] of the fragments in the fragment flow graph is constructed. It turns out that the subtrees of this dominator tree correspond to the required fragment equivalence classes (see Figure 4).

Fragment equivalence classes have many uses [27]. Most importantly, measured isotopomer constraints to fragment F can be directly propagated to another fragment F' in the same equivalence class, by applying the joint carbon mappings between F and F'. This helps in the construction of generalized balance equations (4) where the same isotopomer information is required for each subpool of junction metabolites.

Independence analysis of fragments

A complementary property to fragment equivalence ^{13}C is the statistical independence of fragment isotopomer distributions. Intuitively, if two fragments of the same metabolite are statistically independent, new isotopomer constraints to the union of them can be obtained by taking a product of the isotopomer distributions of the independent fragments.

More formally, the basic question is on what conditions the distribution $D_{M|E\cup F}$ of union of two fragments will necessarily have the form of a *product distribution*

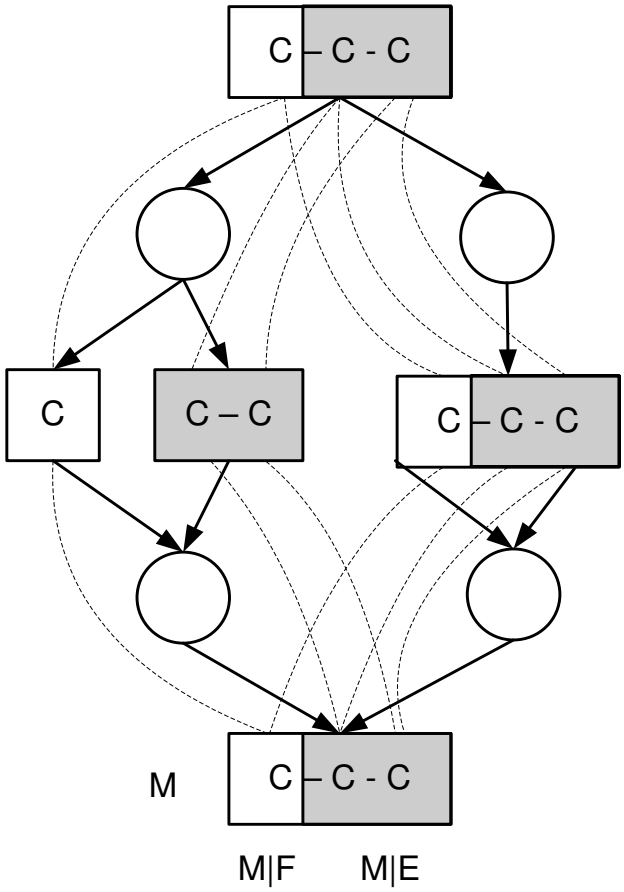

Figure 3
An example of fragment equivalence classes in a branched pathway. An example of equivalence classes of fragments in the metabolic network that contains dominated junction fragments *M*|*E* and *M*|*F*. Grey and white fragments constitute two equivalence classes. Dashed lines depict carbon mappings.

$$D_{M|E\cup F} = D_{M|E} \otimes D_{M|F} \quad (6)$$

where $\otimes$ denotes the tensor product consisting all terms of the form $P_{E\cup F}(b) = P_E(b') \cdot P_F(b'')$, where b' (resp. b'') ranges over all isotopomers of E (resp. F), and b is the isotopomer of $M|E \cup F$ formed by joining E and F.

The utility of fragment independence is in that it gives us constraints to the isotopomer distributions that are complementary to the isotopomer constraints (3) obtained from the measurements.

In general, two criteria need to be satisfied for statistical independence of two fragments $M|E$ and $M|F$. First, the fragments need to be *structurally* independent, meaning that along all pathways producing the metabolite, at some point all carbons of the fragments have resided in different metabolite molecules. This property can be defined in recursive manner. Fragments $M|E$ and $M|F$ are structurally independent if for all carbon pairs (a, b), $a \in E$ and $b \in F$, for all reactions ρ producing M, it holds that

• a and b originate from different reactants of ρ, or

• a and b originate from the same reactant M', and the reactant fragments $M'|F_a$ and $M'|F_b$, where $F_a = \lambda_\rho^{-1}(a)$, $F_b = \lambda_\rho^{-1}(b)$, are structurally independent.

The second necessary condition is that the two fragments need to be dominated by some other metabolite fragments in the network. This will make the fragment distributions flux invariant. Together, the two criteria guarantee (6) to hold.

A simple case of statistical independence of fragments is a (subpool) product metabolite M_i of a single reaction ρ_j, where the fragments $M_i|E$ and $M_i|F$ are disjoint and originate from different reactants. The fragments are structurally independent (by originating from different reactants) and are dominated (by reactant fragments of ρ_j). Hence (6) holds. The underlying assumption here is that enzymes pick their reactants independently and randomly from the available pools. This case of statistical independence of fragments is depicted Figure 1, where white and grey fragments of D-fructose 1,6-biphosphate are statistically independent (in the subpool of reaction ρ_j).

Another simple example is a junction metabolite M_i that has two or more producers with associated subpools M_{ij}. If $M_{ij}|E$ and $M_{ij}|F$ are structurally independent in all subpools, $M_i|E$ and $M_i|F$ are structurally independent as well. If $M_i|E$ and $M_i|F$ are dominated by some fragments in the network, all subpools have the same distribution which takes the form of (6). Without dominance the distribution $D_{M_i|E\cup F}$ will in general be a *flux-dependent mixture* of product distributions $D_{M_i|E\cup F} = \sum_j v_j D_{M_{ij}|E} \otimes D_{M_{ij}|F}$. This case of statistical independence is depicted in Figure 5.

A generalized form of (6), useful for propagation of isotopomer constraints, is derived as follows. Assume independent fragments $M|E$ and $M|F$ of metabolite M and isotopomer constraints $S_{E\cup F}D_{M|E\cup F} = \mathbf{d}_{\mathbf{E}\cup\mathbf{F}}$, $S_E D_{M|E} = \mathbf{d}_\mathbf{E}$ and $S_F D_{M|F} = \mathbf{d}_F$, where $S = S_E \otimes S_F$. Now, the the h'th constraint for fragment E and g'th constraint for fragment F, written as $\Sigma_a s_{ah,E} P_{M|E}(a) = d_{h,E}$ and $\Sigma_c s_{cg,F} P_{M|F}(c) = d_{g,F}$.

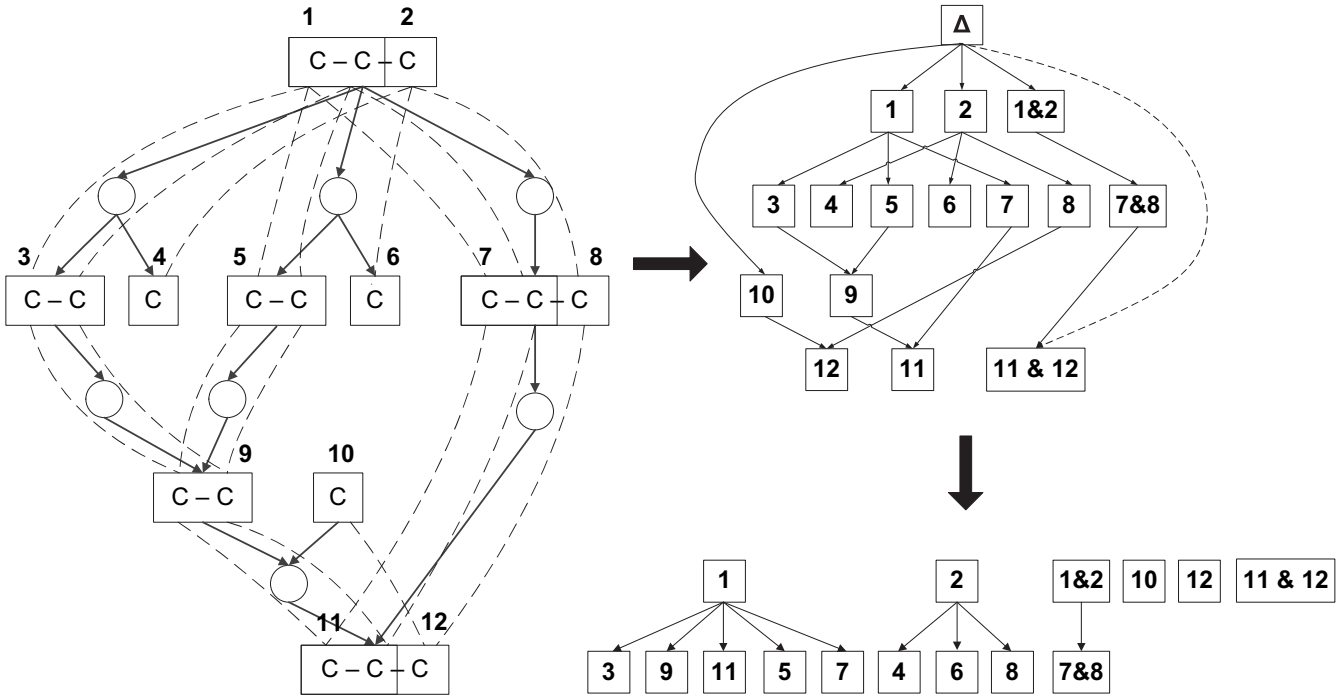

Figure 4
An example of a fragment flow graph and a dominator tree. A metabolic network (left), the corresponding fragment flow graph (up right) and the subtrees of the dominator tree (down right).

Multiplying the constraints, and denoting $s_b = s_{ah,E} \cdot s_{cg,F}$ where b is the isotopomer consistent with fragment isotopomers a and c, we get the following equation

$$d_{h,E} \cdot d_{g,F} = \left(\sum_a s_{ah,E} P_{M|E}(a) \right) \cdot \left(\sum_a s_{cg,F} P_{M|F}(a) \right) = \sum_a s_b P_{M|E\cup F}(b) = d_{l,E\cup F},$$

for the l'th constraint for $E \cup F$. The equations of the above kind can be concisely written in terms of tensors:

$$\mathbf{d} = SD_{M|E\cup F} = (S_E \otimes S_F)\, D_{M|E} \otimes D_{M|F} = \mathbf{d}_E \otimes \mathbf{d}_F. \quad (7)$$

From above, if two of the three vectors $\mathbf{d}$, $\mathbf{d}_E$, $\mathbf{d}_F$ are known, the remaining unknown one can be solved.

We note in passing that computing constraints to the metabolite given constraints to the fragments is straightforward application of (7).

Applying fragment independence analysis to flux ratio computation
In our framework, statistical independence of fragments has two uses. We apply it

1. to compute isotopomer constraints for the union of independent fragments, given isotopomer constraints to its independent fragments, and

2. to compute isotopomer constraints for an independent fragment given isotopomer constraints to the other fragments and the metabolite as a whole.

In both cases making use of (6) gives us a larger set of constraints than the vector space and flow analysis approach alone.

Next we describe how (7) generalizes the basic measurement propagation step of the traditional metabolic flux ratio analysis [31]. In the basic case, the flux ratios are solved for two competing pathways p and q, which p cleaves a certain carbon-carbon bond b of junction M while the q preserves b intact from the external substrate. (See Figure 6 for an example). This serves also as an example of applying (7) to compute isotopomer constraints for the union of independent fragments.

When a uniformly labelled substrate is used, the labelling degree of every carbon in the network is the same (and known *a priori*) when the system reaches isotopomeric steady state. Thus, the isotopomer distribution D_{F_p} of a two-carbon fragment F (metabolite M_F in Figure 6) containing bond b can be computed by (7) for pathway p cleaving b, while for pathway q, D_{F_q} can be propagated

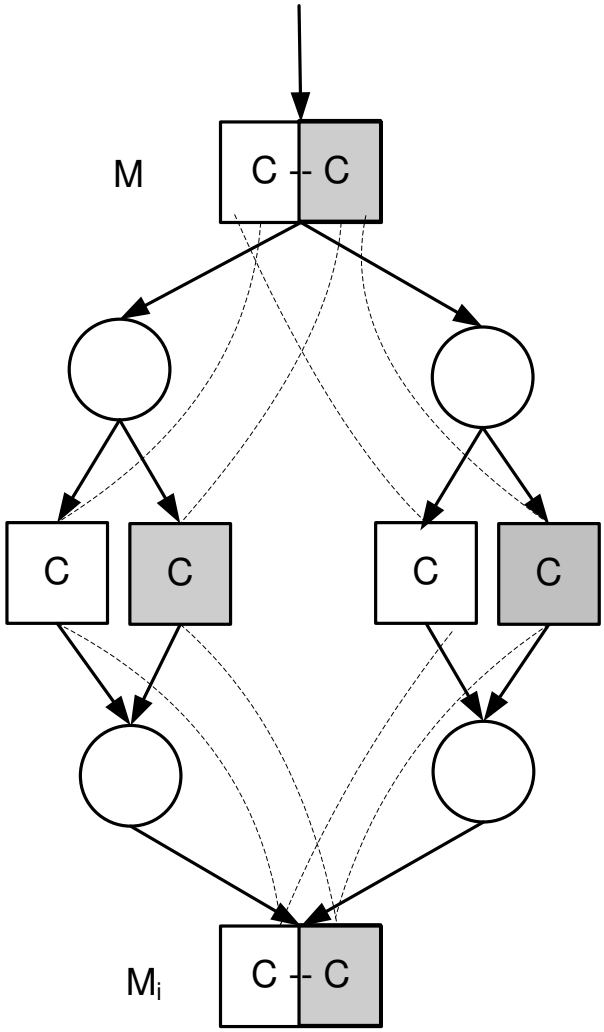

Figure 5
An example of statistical independence of fragments. White and grey one-carbon-fragments of M_i are statistically independent: both fragments are dominated by one-carbon-fragments of *M*, and the fragments are disjoint in every pathway that produce M_i from *M*.

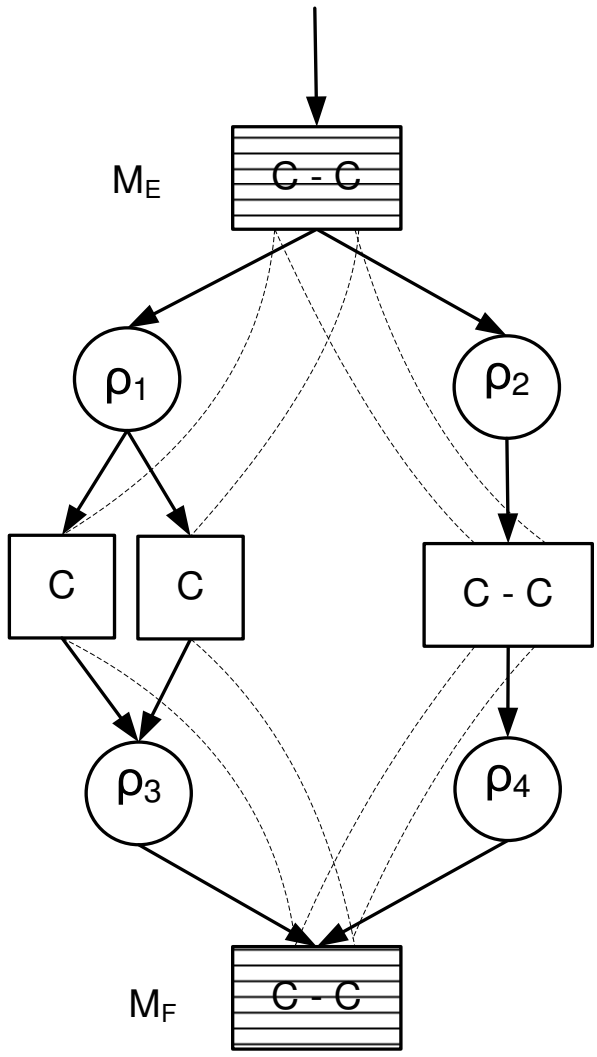

Figure 6
An example of using fragment independence to obtain new isotopomer constraints under uniform substrate labelling. Constraints to the isotopomer distributions of striped metabolites are assumed to be known, either by direct measurement of measurement propagation. In pathway $q = (\rho_2, \rho_4)$, the isotopomer distribution of M_F molecules will be the same as in M_E. In pathway $p = (\rho_1, \rho_3)$, the isotopomer distribution of M_F can be derived by applying fragment independence: the isotopomer distributions of single carbon metabolites produced by ρ_1 are known *a priori* to be equal to the labelling degree of uniformly labelled substrate. As the two carbons of M_{F3} are produced from two different metabolites, these carbons are statistically independent to each other in the subpool and the isotopomer distribution $D(M_{F_p})$ of M_F molecules produced by *p* can be computed by applying Equation 7.

from the external substrate (metabolite M_E in Figure 6) using the fragment equivalence classes of the previous section. If we are able to measure (constraints to) the isotopomer distribution of the mixed pool *F*, we can then automatically derive a generalized isotopomer balance corresponding the manually derived ratio. To use (7) to compute isotopomer constraints for an independent fragment from the known isotopomer constraints to the other independent fragment and to the whole metabolite is complicated by the incompleteness of the measurement data: an arbitrary measurement $SD_{M|E\cup F} = \mathbf{d}$ might not be directly representable via a tensor product $S = S_E \otimes S_F$. Instead, we need to first compute isotopomer subspaces

for known isotopomer constraints where (7) can be applied.

The detailed description of this technique is rather technical and omitted from this article. Here we give an example of the technique (See Figure 7). We assume that we know the mass isotopomer distributions $D^m_{M_i}$ of metabolite M_i (metabolite M_1 in Figure 7) and $D^m_{M_i|E}$ of fragment $M_i|E$. We furthermore assume that $M_i|E$ and $M_i|F$ are independent. From this information the mass isotopomer distribution of $D^m_{M_i|E}$ can be solved. To be exact, $D^m_{M_i|E}$ can be solved from the system containing an equation

$$P_{M_i}(+p) = \sum_{k+l=p} P_E(+k) \cdot P_F(+l) \quad (8)$$

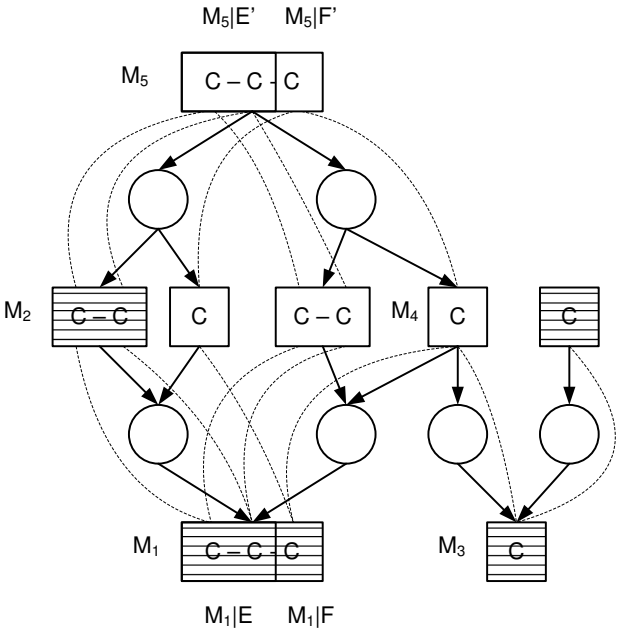

Figure 7
An example of using fragment independence to obtain new isotopomer constraints for a reactant. The mass isotopomer distributions of striped metabolites are assumed to be measured. Fragments $M_1|E$ and M_2 belong to the same fragment equivalence class. Thus, $D_m(M_1|E)$ can be derived from $D_m(M_2)$ by the measurement propagation inside equivalence classes. Furthermore, fragments $M_5|E'$ and $M_5|F'$ dominate fragments $M_1|E$ and $M_1|F$, and the bond between $M_1|E$ and $M_1|F$ is broken in all pathways producing M_1 from M_5. Thus, $M_1|E$ and $M_1|F$ are statistically independent, and $D_m(M_1|F)$ can be deduced from $D_m(M_1)$ and $D_m(M_1|E)$ by utilizing Equation 7. Computed $D_m(M_1|F)$ can then be propagated to M_4, as M_1 and M_4 belong to the same fragment equivalence class. Finally, $D_m(M_4)$ helps to solve the ratios of fluxes entering to M_3.

for each mass isotopomer p of M_i. To see that (8) conforms to (7), we denote the measured mass isotopomer distribution $D^m_{M_i} = S \cdot D(M_i)$ of M_i by $\mathbf{d}_M$ (i.e. rows of coefficient matrix S correspond to different mass isotopomers of M_i) and the measured mass isotopomer distributions of E and F by $\mathbf{d}_E$ and $\mathbf{d}_F$. Let $|E| = 2$, $|F| = 1$ and $|M_i| = 3$, thus $M_i = E \cup F$. We have

$$S_E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, S_F = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

with the tensor product

$$S_E \otimes S_F = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \text{and } S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

As the two matrices are not the same (7) is not directly applicable. However, by summing up the second and the third rows and the fourth and the fifth rows of $S_E \otimes S_F$ we obtain S. Intuitively, this means that Equation (7) can be applied to compute $D^m_{M|F}$, when we take into account that the isotopomer constraints corresponding both second and third rows of $S_E \otimes S_F$ contribute to mass isotopomer $M_i(+1)$, while the isotopomer constraints corresponding fourth and fifth rows of $S_E \otimes S_F$ contribute to mass isotopomer $M_i(+2)$. (From the definition of the tensor product we see that, for example, the second row of $S_E \otimes S_F$ corresponds to isotopomer constraints $P_E(00) \cdot P_F(1) = P_E(+0) \cdot P_F(+1)$ and the third row corresponds to the constraints $P_E(01) \cdot P_F(0) + P_E(10) \cdot P_F(0) = P_E(+1) \cdot P_F(+0)$, thus validating our intuitive observation.) When the similiar information for all rows of $S_E \otimes S_F$ is collected to a linear equation system, we will obtain the following constraints to the mass isotopomer distribution $D^m_{M|F}$ (which in the case of one-carbon-fragment $M|F$ coincides with the isotopomer distribution $D_{M|F}$):

$$\begin{pmatrix} P_E(+0) & 0 \\ P_E(+1) & P_E(+0) \\ P_E(+2) & P_E(+1) \\ 0 & P_E(+2) \end{pmatrix} D_m(M \mid F) = \mathbf{d}_M = \begin{pmatrix} P_{M_i}(+0) \\ P_{M_i}(+1) \\ P_{M_i}(+2) \\ P_{M_i}(+3) \end{pmatrix} = SD_{M_i},$$

which is equal to (8).

Calculability analysis

Isotopomer tracer experiments using less common carbon sources can be very costly because of the prices of purposefully labelled substrates. Thus, it is very useful to be able to first conduct *in silico* calculability analysis to find out, whether it is even in principle possible to obtain the required flux information from the tracer experiment. By analyzing the fragment equivalence classes, it is relatively easy to perform this kind of "structural identifiability analysis" (cf. [39,40] for global isotopomer balancing), that is, to discover the set of junction metabolites for which the flux ratios can be calculated (in the best case) from the given measurement data: it is enough to check what type of isotopomer constraints

$$\sum_b s_{bij} P_{M_{ij}}(b) = d_{ij}, \qquad (9)$$

can be propagated to each subpool M_{ij} of junction metabolites M_i from the measured metabolites (we need to know only coefficients s_{bij}, not the isotopomer abundances d_{ij}). Then, by applying the techniques of computing vector subspace intersection described above, we can compute the maximal number of linearly independent constraints obtainable for the flux ratios of each junction. Thus, it is possible to check before costly and time-consuming wet lab experiments, whether the experiments even have potential to answer the biological questions at hand. The results of the calculability analysis tell which flux ratios are in principle determinable, given the labelling of external substrates, topology of the metabolic network and the available measurement data. It then depends on the actual flux distribution and the accuracy of the measurements, whether these ratios can be reliably determined from the experimental data.

Estimating the flux distribution of the metabolic network

In the main step of our framework for ^{13}C metabolic flux analysis, the fluxes of the metabolic network are estimated by forming and solving generalized isotopomer balance equations (4). The generalized isotopomer balance equations are based on the isotopomer measurement data that is first propagated in the network to unmeasured metabolites by utilizing the results of the structural analysis presented above.

Measurement propagation

The aim of the propagation of measurement data is to infer from the isotopomer constraints of measured metabolites as many isotopomer constraints as possible to unmeasured metabolites. As a rule of thumb, more constraints the unmeasured metabolites will get more generalized isotopomer balance equations (4) bounding the fluxes can be written.

Fragment equivalence classes can be utilized in the measurement propagation: from isotopomer constraints known for fragment $M_i|F$ isotopomer constraints for other fragments $M_l|F_k$ in the equivalence class of F can be easily computed. The process is the following:

1. Before measurements are propagated from fragment $M|F$ of measured metabolite M to other fragments in the equivalence class of F, isotopomer constraints to F are computed from the constraints measured to the whole metabolite M by using the vector space projection techniques (see Section *Projection of isotopomer measurements to fragments*).

2. The fragment constraints are propagated to all fragments F' that have been found equivalent to F via the flow analysis technique.

This requires mapping of isotopomers of F to isotopomers of F' by applying the carbon mappings of the reactions along any pathway between F and F'.

3. After the propagation of measurement data inside the fragment equivalence classes, new isotopomer constraints for independent fragments of the same metabolite can be derived, as described in Section *Independence analysis of fragments*.

Steps 2 and 3 can be iterated until no new isotopomer constraints to the fragments are discovered.

Construction of generalized isotopomer balances

After the propagation step, we have some isotopomer constraints $S_{ij}D_{M_{ij}} = \mathsf{d}_{ij}$ for each subpool j of every junction metabolite M_i. (For non-junction metabolites, isotopomer balance equations do not contain any additional flux information compared to the mass balances.) In the best case we know complete isotopomer distribution $D_{M_{ij}}$, in the worst case we have only trivial constraints stating that the sum of relative abundances of all isotopomers equals one.

Next, a linear equation system containing flux constraints obtained from mass balances (1) and generalized isotopomer balances (4) is constructed.

However, the isotopomer constraints of different subpools do not yet conform to (4) as the matrices S_{ij} are not necessarily the same.

Thus we still need to compute a common subspace $\mathcal{Y}_i = \bigcap_j \mathcal{S}_{ij}$ ($\mathcal{S}_{ij}$ is spanned by the rows of S_{ij}) of the isotopomer constraints known for each subpool M_{ij} and project subpool constraints $S_{ij}D_{M_{ij}} = \mathsf{d}_{ij}$ to $\mathcal{Y}_i$.

This can be done with the same techniques that were previously applied to project measured isotopomer information of a metabolite to its fragments. Let Y_i be a matrix with row space $\mathcal{Y}_i$. After the projection we obtain isotopomer constraints $Y_i D_{M_{ij}} = \mathsf{z}_{ij}$ for each subpool M_{ij} (See Figure 8 for an example).

Now the isotopomer constraints of all the subpools lie in the same subspace of $\mathcal{I}_{M_i}$ and we are ready to write the system of generalized isotopomer balance equations (4) for every junction M_i:

$$\sum_{j=1}^{n} \alpha_{ij} v_j \mathsf{z}_{ij} = \beta_i \mathsf{z}_{i0}, \qquad (10)$$

that is,

$$A_i \mathsf{v} = \begin{bmatrix} \alpha_{1i}(\mathsf{z}_{1i})_1 & \cdots & \alpha_{ni}(\mathsf{z}_{ni})_1 \\ \vdots & \ddots & \vdots \\ \alpha_{1i}(\mathsf{z}_{1i})_r & \cdots & \alpha_{ni}(\mathsf{z}_{ni})_r \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \mathsf{g}_i, \qquad (11)$$

where $\mathbf{g}_i = \beta_i \mathbf{z}_{i0}$.

Estimating the fluxes

The ratios of (forward) fluxes producing M_i can be computed by solving the corresponding Equation (11) augmented with a constraint that fixes the out flow from M_i to equal 1. Thus, we obtain flux ratios of junction metabolites without manual derivation of ratio equations, without nonlinear optimization and without knowing intake and outtake rates of external metabolites or biomass composition.

In addition, when the equations (11) of all junction metabolites are combined with the mass balances (1) of non-junctions, we obtain a linear equation system

$$A\mathsf{v} = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} \cdot \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \mathsf{g}_1 \\ \vdots \\ \mathsf{g}_m \end{bmatrix} = \mathsf{g} \qquad (12)$$

constraining the fluxes **v** of the network that contains a block (junctions) or a row (non-junctions) A_k for each metabolite M_k. Measured external fluxes and other known constraints, such as the composition of biomass can also

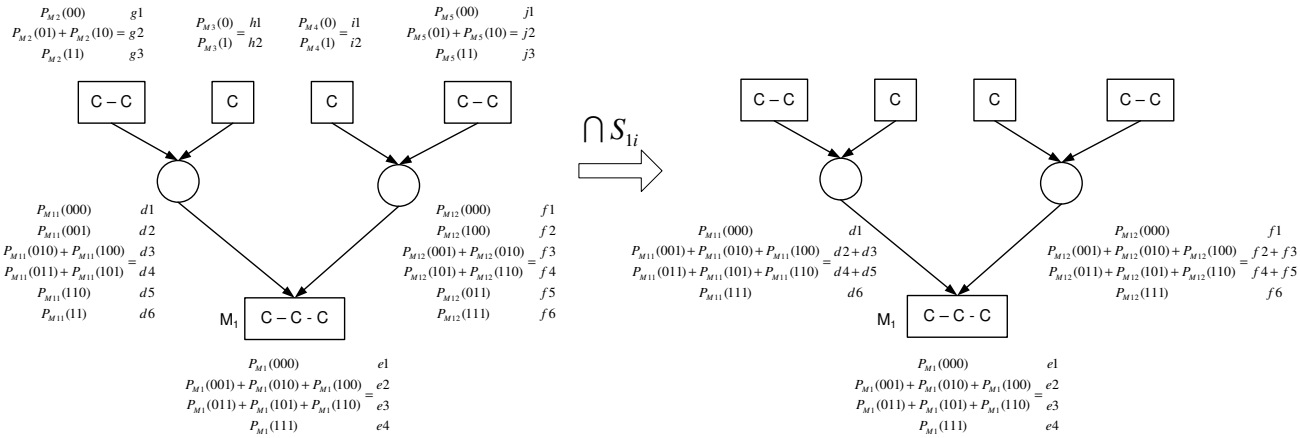

Figure 8

An example of the computation of the common subspace of isotopomer constraints in different subpools. The mass isotopomer distribution of junction metabolite M_1 is assumed to be measured. For the in flow subpools M_{11} and M_{12} we obtain isotopomer constraints from the above reactant metabolites by measurement propagation. These propagated constraints must be projected to mass isotopomer to the subspace defined by the mass isotopomer distribution of M_1 before generalized isotopomer balances are constructed.

be added to (12) as additional constraints. Additional constraints, like ones derived from gene regulatory information [41] or from thermodynamic analysis of metabolism [42-44] can also easily be included to (12).

If (12) is of full rank, the whole flux distribution can be solved with standard linear algebra [45]. Also, more complex, nonlinear methods can be applied to model the effect of experimental errors to the estimated flux distribution [20]. In a common case where the system is of less than full rank, a single flux distribution can not be pinpointed without additional constraints. Instead, (12) defines the space of feasible flux distributions, that are all equally good solutions. In that case we can apply techniques developed for the analysis of stoichiometric matrices to determine as many fluxes as possible [46] from (12). More generally, by linear programming we can obtain maximum (resp. minimum) values for each flux v_i:

$$\begin{aligned} &\text{For each } v_i: \\ &\qquad \max \quad v_i \\ &\qquad \text{s.t.} \quad A\mathbf{v} = \mathbf{g} \\ &\qquad v_i^{min} \le v_i \le v_i^{max} \forall v_i \in \mathbf{v}, \end{aligned} \tag{13}$$

where v_i^{min} and v_i^{max} are predetermined minimum and maximum allowable values for v_i

Furthermore, it is possible to search for in some sense optimal flux distribution – for example a flux distribution maximizing the production of biomass – from the feasible space defined by (12) by linear programming techniques of flux balance analysis [1,3,47,48]. In that case, isotopomer data constrain the feasible space more than the stoichiometric information would alone do, thus possibly allowing more accurate estimations of the real flux distribution.

Statistical analysis

For an experimentalist, it is important to know how sensitive the obtained estimation of fluxes is to measurement errors. If enough repeated measurements are not available to assess this sensitivity, it has to be estimated by computational techniques. In the global isotopomer balancing framework for ^{13}C metabolic flux analysis, many mathematically or computationally involved methods have been developed to analyze the sensitivity of estimated flux distributions to errors in isotopomer measurements and the sensitivity of the objective function to the changes in the generated candidate flux distributions [49-53].

As our direct method for ^{13}C metabolic flux analysis is computationally efficient, we can afford to a simple, yet powerful Monte Carlo procedure to obtain estimates on the variability of individual fluxes due to measurement errors:

1. For each measured metabolite M_i: By studying the variability in the repeated measurements, fix the distribution Ω_i from which the measurements of M_i are sampled.

2. Repeat k times:

(a) For each measured metabolite M_i: sample a measurement from Ω_i.

(b) Estimate fluxes $\mathbf{v}_l$ from the sampled measurements.

3. Compute appropriate statistics from the set $V = \{\mathbf{v}_1, ..., \mathbf{v}_k\}$ to describe the sensitivity of fluxes to measurement errors.

Possible statistics that can be applied in the last step of the above algorithm include standard deviation, empirical confidence intervals [53], kurtosis, standard error etc. of each individual flux v_j and measures of "compactness" of V, such as (normalized) average distance of items of V from the sample average.

Experimental NMR and GC-MS methods

In this section we shortly describe the experimental procedures applied in NMR and GC-MS isotopomer measurements that produced the data for Section.

In the first experiment *S. cerevisiae* was grown in an aerobic glucose-limited chemostat culture at dilution rate 0.1 h^{-1}. After reaching a metabolic steady state, as determined by constant physiological parameters 10% of the carbon source in the medium was replaced with fully carbon labelled glucose ([U-13C]) for approximately 1.5 residence times, so that about 78% of the biomass was uniformly labelled. 2D [^{13}C, ^{1}H] COSY spectra of harvested and hydrolysed biomass were acquired for both aliphatic and aromatic resonances at 40°C on a Varian Inova 600 MHz NMR spectrometer. The software FCAL v.2.3.0 [19] was used to compute isotopomer constraints for 15 amino acids from the spectra. Detailed description of the cultivation set up can be found in [54] whereas similar ^{13}C labeling set up, NMR experiments and spectral data analysis as were applied here have been described in [55].

In the second experiment *B. subtilis* was grown on shake flasks containing 50 ml M9 minimal medium. In the experiment, the medium was supplemented with 50 mg/L tryptophan and 3 g/L glucose labelled to the first carbon position ([1-13C]) (99%; Cambridge Isotope Laboratories) or a mixture of 0.6 g/L fully carbon labelled glucose ([U-13C]) (99%; Cambridge Isotope Laboratories) and 2.4 g/L unlabeled glucose as the sole carbon source. Four-

teen derivatized amino acids were analyzed for ^{13}C labeling patterns with a series 8000GC combined with an MD800 mass spectrometer (Fisons instruments). More information about the details of the measurement procedure can be found from [20] where identical measurement techniques were applied.

Authors' contributions

AR, JR and EU are the main contributors of the computational methods given in the article. AR implemented the methods and ran the experiments. PJ constructed the models of metabolic networks and contributed to the development of the computational methods. NZ and HM contributed the development of the computational methods. AR, JR, NZ and PJ wrote the manuscript. All authors have read and accepted the contents of the manuscript.

Additional material

Additional file 1

Model of central carbon metabolism of* S. cerevisiae *in SBML format. *Additional file 1 contains the stoichiometric model of central carbon metabolism of* S. cerevisiae *used in the experiments of the article. The model is provided as a text file containing SBML compliant descriptions of metabolites and reactions in the model. Metabolites are identified by their KEGG LIGAND codes.*

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2105-9-266-S1.txt]

Additional file 2

***Illustration of the model of central carbon metabolism of* S. cerevisiae.** *Bidirectional reactions in PPP pathways are depicted as unidirectional for better readability.*

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2105-9-266-S2.pdf]

Additional file 3

Model of central carbon metabolism of* B. subtilis *in SBML format. *Additional file 2 contains the stoichiometric model of central carbon metabolism of* B. subtilis *used in the experiments of the article. The model is provided as a text file containing SBML compliant descriptions of metabolites and reactions in the model. Metabolites are identified by their KEGG LIGAND codes.*

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2105-9-266-S3.txt]

Additional file 4

***Illustration of the model of central carbon metabolism of* B. subtilis.** *Bidirectional reactions in PPP pathways are depicted as unidirectional for better readability.*

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2105-9-266-S4.pdf]

Acknowledgements

This research has been supported by Academy of Finland (SYSBIO programme, grant 207436, SYSFYS project). We thank Uwe Sauer and Merja Penttilä for providing the data and for their comments to the work presented in the article; Markus Heinonen, Esa Pitkänen and Arto Åkerlund for the implementation of software tools supporting the presented framework; Roelco Kleijn, Sarah-Maria Fendt, Simon Tännler and Martin Rühl for providing the data and for fruitful discussions. We also thank the anonymous reviewers whose comments helped us to improve the manuscript substantially.

References

1. Varma A, Palsson B: **Metabolic flux balancing: basic concepts, scientific and practical use.** *Nature Biotechnology* 1994, **12(10):**994-998.
2. Edwards JS, Ibarra RU, Palsson B: **In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data.** *Nature Biotechnology* 2001, **19(2):**125-130.
3. Schütz R, Kuepfer L, Sauer U: **Systematic evaluation of objective functions for predicting intracellular fluxes in *E. coli*.** *Molecular Systems Biology* 2007, **3(119):**.
4. Szyperski T: **Biosynthetically directed fractional ^{13}C-labelling of proteinogenic amino acids.** *European Journal of Biochemistry* 1995, **232(2):**433-448.
5. Dauner M, Sauer U: **GC-MS analysis of amino acids rapidly provides rich information for isotopomer balancing.** *Biotechnology Progress* 2000, **16(4):**642-649.
6. Rousu J, Rantanen A, Ketola R, Kokkonen J: **Isotopomer distribution computation from tandem mass spectrometric data with overlapping fragment spectra.** *Spectroscopy* 2005, **19:**53-67.
7. Schmidt K, Carlsen M, Nielsen J, Viladsen J: **Modeling Isotopomer Distributions in Biochemical Networks Using Isotopomer Mapping Matrices.** *Biotechnology and Bioengineering* 1997, **55(6):**831-840.
8. Wiechert W, Petersen Möllney S, de Graaf A: **A Universal Framework for ^{13}C Metabolic Flux Analysis.** *Metabolic Engineering* 2001, **3(3):**265-283.
9. Wiechert W, de Graaf A: **Bidirectional Reaction Steps in Metabolic Networks: I. Modeling and simulation of carbon isotope labeling experiments.** *Biotechnology and Bioengineering* 1997, **55(1):**101-117.
10. Antoniewicz M, Kelleher J, Stephanopoulos G: **Elementary Metabolite Units (EMU): a novel framework for modeling isotopic distributions.** *Metabolic Engineering* 2007, **9(1):**68-86.
11. van Winden W, Heijnen J, Verheijen P: **Cumulative bondomers: a new concept in flux analysis from 2D [^{13}C, ^{1}H] COSY NMR data.** *Biotechnology and Bioengineering* 2002, **80(7):**731-745.
12. Nöh K, Wahl A, Wiechert W: **Computational tools for isotopically instationary ^{13}C labeling experiments under metabolic steady state conditions.** *Metabolic Engineering* 2006, **8(6):**554-577.
13. Nöh K, Grönke K, Luo B, Takors R, Oldiges M, Wiechert W: **Metabolic flux analysis at ultra short time scale: Isotopically nonstationary ^{13}C labeling experiments.** *Journal of Biotechnology* 2007, **129(2):**249-267.
14. Young J, Walther J, M A, Yoo H, Stephanopoulos G: **An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis.** *Biotechnology and Bioengineering* 2007, **99(3):**686-699.
15. Ghosh S, Zhu T, Grossmann I, Ataai M, Domach M: **Closing the loop between feasible flux scenario identification for construct evaluation and resolution of realized fluxes via NMR.** *Computers and Chemical Engineering* 2005, **29(3):**459-466.
16. Blank L, Kuepfer L, Sauer U: **Large-scale ^{13}C-flux analysis reveals mechanistic principles metabolic network robustness to null mutations in yeast.** *Genome Biology* 2005, **6(6):**R49.
17. Blank L, Lehmbeck F, Sauer U: **Metabolic-flux and network analysis in fourteen hemiascomycetous yeasts.** *FEMS Yeast Research* 2005, **5(6–7):**545-558.
18. Fischer E, Sauer U: **Metabolic flux profiling of *Escherichia coli* mutants in central carbon metabolism using GC-MS.** *European Journal of Biochemistry* 2003, **270(5):**880-891.

19. Szyperski T, Glaser R, Hochuli M, Fiaux J, Sauer U, Bailey J, Wütrich K: **Bioreaction Network Topology and Metabolic Flux Ratio Analysis by Biosynthetic Fractional ^{13}C Labeling and Two-Dimensional NMR Spectrometry.** *Metabolic Engineering* 1999, **1(3):**189-197.
20. Fischer E, Zamboni N, Sauer U: **High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived ^{13}C constraints.** *Analytical Biochemistry* 2004, **325(2):**308-316.
21. Rousu J, Rantanen A, Maaheimo H, Pitkänen E, Saarela K, Ukkonen E: **A method for estimating metabolic fluxes from incomplete isotopomer information.** *Computational Methods in Systems Biology, Proceedings of the First International Workshop, CMSB Volume 2602 of Lecture Notes in Computer Science* 2003:88-103.
22. Sauer U, Hatzimanikatis V, Bailey J, Hochuli M, Szyperski T, Wüthrich K: **Metabolic fluxes in riboflavin-producing *Bacillus subtilis*.** *Nature Biotechnology* 1997, **15(5):**448-452.
23. Maaheimo H, Fiaux J, Cakar Z, Bailey J, Sauer U, Szyperski T: **Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional ^{13}C labeling of common amino acids.** *European Journal of Biochemistry* 2001, **268(8):**2464-2479.
24. Zamboni N, Sauer U: **Knockout of the high-coupling cytochrome aa3 oxidase reduces TCA cycle fluxes in *Bacillus subtilis*.** *FEMS Microbiology Letters* 2003, **226(1):**121-126.
25. Sola A, Maaheimo H, Ylönen K, Ferrer P, Szyperski T: **Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*.** *European Journal of Biochemistry* 2004, **271(12):**2462-2470.
26. Sola A, Jouhten P, Maaheimo H, Sanchez-Ferrando F, Szyperski T, Ferrer P: **Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates.** *Microbiology* 2007, **153(1):**281-290.
27. Rantanen A, Maaheimo H, Pitkänen E, Rousu J, Ukkonen E: **Equivalence of metabolite fragments and flow analysis of isotopomer distributions for flux estimation.** *Transactions on Computational Systems Biology VI, Lecture Notes in Bioinformatics* 2006, **4220:**198-220.
28. Kleijn R, Geertman J, Nfor B, Ras C, Schipper D, Pronk J, Heijnen J, van Maris A, van Winden W: **Metabolic flux analysis of a glycerol-overproducing *Saccharomyces cerevisiae* strain based on GC-MS, LC-MS and NMR derived ^{13}C-labeling data.** *FEMS Yeast Research* 2006, **7(2):**216-231.
29. Arita M: **In Silico Atomic Tracing of Substrate-Product Relationships in *Escherichia coli* Intermediary Metabolism.** *Genome Research* 2003, **13(11):**2455-2466.
30. Zamboni N, Fischer E, Sauer U: **FiatFlux – a software for metabolic flux analysis from ^{13}C-glucose experiments.** *BMC Bioinformatics* 2005, **6(209):**.
31. Szyperski T: **^{13} C-NMR, MS and metabolic flux balancing in biotechnology research.** *Quarterly Reviews of Biophysics* 1998, **31(1):**41-106.
32. Weitzel M, Wiechert W, Nöh K: **The topology of metabolic isotope labeling networks.** *BMC Bioinformatics* 2007, **8(315):**.
33. van Winden W, van Dam J, Ras C, Kleijn R, Vinke J, Gulik W, Heijnen J: **Metabolic-flux analysis of *Saccharomyces cerevisiae* CEN.PK113-7D based on mass isotopomer measurements of ^{13}C-labeled primary metabolites.** *FEMS Yeast Research* 2005, **5(6–7):**559-568.
34. Rantanen A, Mielikäinen T, Rousu J, Maaheimo H, Ukkonen E: **Planning optimal measurements of isotopomer distributions for estimation of metabolic fluxes.** *Bioinformatics* 2006, **22(10):**1198-1206.
35. Schmidt K, Nielsen J, Villadsen J: **Quantitative analysis of metabolic fluxes in *Escherichia coli*, using two-dimensional NMR spectroscopy and complete isotopomer models.** *Journal of Biotechnology* 1999, **71(1–3):**175-189.
36. Rantanen A, Rousu J, Ketola R, Kokkonen J, Tarkiainen V: **Computing Positional Isotopomer Distributions from Tandem Mass Spectrometric Data.** *Metabolic Engineering* 2002, **4:**285-294.
37. Aho A, Hopcroft J, Ullman J: *The Design and Analysis of Computer Algorithms* Addison Wesley; 1974.
38. Appel A: *Modern Compiler Implementation in Java* Cambridge University Press; 1998.
39. Isermann N, Wiechert W: **Metabolic isotopomer labeling systems. Part II: structural identifibiality analysis.** *Mathematical Biosciences* 2003, **183(2):**175-214.
40. van Winden W, Heijnen J, Verheijen P, Grievink J: **A priori analysis of metabolic flux identifiability from ^{13}C-labeling data.** *Biotechnology and Bioengineering* 2001, **74(6):**505-516.
41. Covert M, Schilling C, Palsson B: **Regulation of Gene Expression in Flux Balance Models of Metabolism.** *Journal of Theoretical Biology* 2001, **213(1):**73-88.
42. Kümmel A, Panke S, Heinemann M: **Systematic assignment of thermodynamic constraints in metabolic network models.** *BMC Bioinformatics* 2006, **7(512):**.
43. Henry C, Broadbelt L, Hatzimanikatis V: **Thermodynamics-Based Metabolic Flux Analysis.** *Biophysical Journal* 2007, **92(5):**1792-1805.
44. Hoppe A, Hoffmann S, Holzhütter H: **Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks.** *BMC Systems Biology* 2007, **1(23):**.
45. Schwarz H: *Numerical Analysis: A Comprehensive Introduction* John Wiley & Sons; 1989.
46. Klamt S, Schuster S: **Calculating as many fluxes as possible in underdetermined metabolic networks.** *Molecular Biology Reports* 2002, **29(1–2):**243-248.
47. Bonarius H, Schmidt G, Tramper J: **Flux analysis of underdetermined metabolic networks: the quest for the missing constraints.** *Trends in Biotechnology* 1997, **15(8):**308-314.
48. Edwards J, Covert M, Palsson B: **Metabolic modelling of microbes: the flux-balance approach.** *Environmental Microbiology* 2002, **4(3):**133-140.
49. Möllney M, Wiechert W, Kownatzki D, de Graaf A: **Bidirectional Reaction Steps in Metabolic Networks IV: Optimal Design of Isotopomer Labeling Experiments.** *Biotechnology and Bioengineering* 1999, **66(2):**86-103.
50. Kleijn R, van Winden W, Ras C, van Gulik W, Schipper D, Heijnen J: **13C-Labeled Gluconate Tracing as a Direct and Accurate Method for Determining the Pentose Phosphate Pathway Split Ratio in *Penicillium chrysogenum*.** *Applied and Environmental Microbiology* 2006, **72(7):**4743-4754.
51. Araúzo-Bravo M, Shimizu K: **An improved method for statistical analysis of metabolic flux analysis using isotopomer mapping matrices with analytical expressions.** *Journal of Biotechnology* 2003, **105(1–2):**117-133.
52. Wiechert W, Siefke C, de Graaf A, Marx A: **Bidirectional Reaction Steps in Metabolic Networks: II. Flux Estimation and Statistical Analysis.** *Biotechnology and Bioengineering* 1997, **55(1):**118-134.
53. Antoniewicz M, Kelleher J, Stephanopoulos G: **Determination of confidence intervals of metabolic fluxes estimated from stable isotopome measurements.** *Metabolic Engineering* 2006, **8(4):**324-337.
54. Wiebe M, Rintala E, Tamminen A, Simolin H, Salusjärvi L, Toivari M, Kokkonen J, Kiuru J, Ketola R, Jouhten P, Huuskonen A, Maaheimo H, Ruohonen L, Penttilä M: **Central carbon metabolism of *Saccharomyces cerevisiae* in anaerobic, oxygen-limited and fully aerobic steady-state conditions and following a shift to anaerobic conditions.** *FEMS Yeast Research* 2008, **8(1):**140-154.
55. Fiaux J, Cakar P, Sonderegger M, Wüthrich K, Szyperski T, Sauer U: **Metabolic-Flux Profiling of the Yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*.** *Eukaryotic Cell* 2003, **2(1):**170-180.

Series title, number and report code of publication

VTT Publications 724
VTT-PUBS-724

Author(s)
Paula Jouhten

Title

Metabolic modelling and ^{13}C flux analysis.
Application to biotechnologically important yeasts and a fungus

Abstract

All bioconversions in cells derive from metabolism. Microbial metabolisms contain potential for bioconversions from simple source molecules to unlimited number of biochemicals and for degradation of even detrimental compounds. Metabolic fluxes are rates of consumption and production of compounds in metabolic reactions. Fluxes emerge as an ultimate phenotype of an organism from an integrated regulatory function of the underlying networks of complex and dynamic biochemical interactions. Since the fluxes are time-dependent, they have to be inferred from other, measurable, quantities by modelling and computational analysis. ^{13}C-labelling is crucial for quantitative analysis of fluxes through intracellular alternative pathways. Local flux ratio analysis utilises uniform ^{13}C-labelling experiments, where the carbon source contains a fraction of uniformly ^{13}C-labelled molecules. Carbon-carbon bonds are cleaved and formed in metabolic reactions depending on the *in vivo* fluxes. ^{13}C-labelling patterns of metabolites or macromolecule components can be detected by mass spectrometry (MS) or nuclear magnetic resonance (NMR) spectroscopy. Local flux ratio analysis utilises directly the ^{13}C-labelling data and metabolic network models to solve ratios of converging fluxes.

In this thesis the local flux ratio analysis has been extended and applied to analysis of phenotypes of biotechnologically important yeasts *Saccharomyces cerevisiae* and *Pichia pastoris,* and a fungus *Trichoderma reesei*. Oxygen dependence of *in vivo* net flux distribution of *S. cerevisiae* was quantified by using local flux ratios as additional constraints to the stoichiometric model of the central carbon metabolism. The distribution of fluxes in the pyruvate branching point turned out to be most responsive to different oxygen availabilities. The distribution of fluxes was observed to vary not only between the fully respiratory, respiro-fermentative and fermentative metabolic states but also between different respiro-fermentative states. The local flux ratio analysis was extended to the case of two-carbon source of glycerol and methanol co-utilisation by *P. pastoris*. The fraction of methanol in the carbon source did not have as profound effect on the distribution of fluxes as the growth rate. The effect of carbon catabolite repression (CCR) on fluxes of *T. reesei* was studied by reconstructing amino acid biosynthetic pathways and by performing local flux ratio analysis. *T. reesei* was observed to primarily utilise respiratory metabolism also in conditions of CCR. *T. reesei* metabolism was further studied and L-threo-3-deoxy-hexulosonate was identified as L-galactonate dehydratase reaction product by using NMR spectroscopy. L-galactonate dehydratase reaction is part of the fungal pathway for D-galacturonic acid catabolism.

ISBN
978-951-38-7371-4 (soft back ed.)
978-951-38-7372-1 (URL: http://www.vtt.fi/publications/index.jsp)

Series title and ISSN	Project number
VTT Publications 1235-0621 (soft back ed.) 1455-0849 (URL: http://www.vtt.fi/publications/index.jsp)	

Date	Language	Pages
November 2009	English, Finnish abstr.	94 p. + app. 83 p.

Name of project	Commissioned by

Keywords	Publisher
Metabolic modelling, metabolic flux, metabolic flux analysis (MFA), ^{13}C-labelling, ^{13}C-MFA, nuclear magnetic resonance (NMR) spectroscopy	VTT Technical Research Centre of Finland P.O.Box 1000, FI-02044 VTT, Finland Phone internat. +358 20 722 4520 Fax +358 20 722 4374

Julkaisun sarja, numero ja raporttikoodi

VTT Publications 724
VTT-PUBS-724

Tekijä(t)
Paula Jouhten

Nimeke

Aineenvaihdunnan mallinnus ja ^{13}C-vuoanalyysi
Sovellukset bioteknologisesti tärkeisiin hiivoihin ja homeeseen

Tiivistelmä

Aineenvaihdunta kattaa kaikki biomuunnokset soluissa. Mikrobiaineenvaihdunta mahdollistaa yksinkertaisten lähtöaineiden muuntamisen rajoittamattomaksi määräksi erilaisia biokemikaaleja ja jopa haitallisten aineiden hajottamisen. Aineenvaihduntavuot ovat yhdisteiden kulutus- ja tuottonopeuksia aineenvaihdunnan reaktioissa. Vuot ilmentyvät organismin todellisena fenotyyppinä, jota säätelevät yhteistoiminnallisesti solun monimutkaiset ja dynaamiset vuorovaikutusverkot. Koska vuot ovat aikariippuvaisia, ne on määritettävä mallinnuksen ja laskennallisen analyysin avulla toisista, mitattavissa olevista, suureista. ^{13}C-leimaus on välttämätöntä, jotta vuot vaihtoehtoisilla solunsisäisillä reiteillä voidaan määrittää kvantitatiivisesti. Paikallisessa vuosuhdeanalyysissa käytetään tasaista ^{13}C-leimausta, jossa hiilenlähde sisältää osuuden täydellisesti ^{13}C-leimattuja molekyylejä. *In vivo* -vuot määräävät missä suhteissa aineenvaihdunnassa katkeaa ja muodostuu uusia hiili-hiilisidoksia. Aineenvaihdunnan välituotteiden ja makromolekyylien komponenttien ^{13}C-leimauskuvioita voidaan mitata massaspektrometrialla (MS) tai ydinmagneettisella resonanssispektroskopialla (NMR). Paikallisessa vuosuhdeanalyysissa käytetään suoraan mittausinformaatiota ^{13}C-leimauskuvioista ja aineenvaihduntaverkkomalleja vuosuhteiden ratkaisemiseksi.

Väitöskirjassa paikallista vuosuhdeanalyysia laajennettiin ja sovellettiin bioteknologisesti tärkeiden hiivojen *Saccharomyces cerevisiae* ja *Pichia pastoris,* ja homeen *Trichoderma reesei* fenotyyppien analysoimiseksi. *S. cerevisiae*n *in vivo* -vuojakauman riippuvuus hapen saatavuudesta määritettiin kvantitatiivisesti käyttämällä paikallisia vuosuhteita lisärajoitteina keskeisen hiiliaineenvaihdunnan stoikiometriselle mallille. Pyruvaattiristeyksen vuojakauma osoittautui herkimmäksi eri happisaatavuuksille. Selvästi erilaiset vuojakaumat havaittiin täysin respiratiivisessa, respiro-fermentatiivisessa ja täysin fermentatiivisessa aineenvaihdunnan tilassa, mutta myös eri respiro-fermentatiivisissa tiloissa. Paikallinen vuosuhdeanalyysi laajennettiin kahden hiilenlähteen tapaukseen, jossa *P. pastoris* kulutti samanaikaisesti glyserolia ja metanolia. Metanolin osuudella kokonaishiilenlähteessä ei ollut yhtä merkittävää vaikutusta vuojakaumaan kuin hiivan kasvunopeudella. Hiilikataboliittirepression (CCR) vaikutusta *T. reesei*n vuojakaumaan tutkittiin rekonstruoimalla aminohapposynteesireitit ja tekemällä paikallinen vuosuhdeanalyysi. *T. reesei*n havaittiin käyttävän pääasiassa respiratiivista aineenvaihduntaa myös repressoivissa olosuhteissa. NMR-spektroskopiaa käytettiin myös D-galakturonihapon kaboliareitin tutkimuksessa ja L-treo-3-deoksi-heksulonaatti tunnistettiin *T. reesei*n L-galaktonaattidehydrataasireaktion tuotteeksi.

ISBN
978-951-38-7371-4 (nid.)
978-951-38-7372-1 (URL: http://www.vtt.fi/publications/index.jsp)

Avainnimeke ja ISSN	Projektinumero
VTT Publications 1235-0621 (nid.) 1455-0849 (URL: http://www.vtt.fi/publications/index.jsp)	

Julkaisuaika	Kieli	Sivuja
Marraskuu 2009	Englanti, suom. tiiv.	94 s. + liitt. 83 s.

Projektin nimi	Toimeksiantaja(t)

Avainsanat	Julkaisija
Metabolic modelling, metabolic flux, metabolic flux analysis (MFA), ^{13}C-labelling, ^{13}C-MFA, nuclear magnetic resonance (NMR) spectroscopy	VTT PL 1000, 02044 VTT Puh. 020 722 4520 Faksi 020 722 4374

Metabolisms of microorganisms contain possibilities for conversions of simple source molecules to unlimited number of biochemicals and for degradation of even hazardous compounds. Rates of metabolic reactions are called fluxes. They are in sense process streams of a cell factory in case of biotechnologically important organisms. Since the fluxes are time-dependent, they cannot be directly measured but have to be inferred from other, measurable, quantities by modelling and computational analysis. ^{13}C-labelling is crucial for quantitative analysis of fluxes through alternative pathways inside the cells. Fluxes emerge as an ultimate phenotype of an organism from an integrated regulatory function of the underlying networks of complex and dynamic biochemical interactions. Inferring fluxes and their regulation in simple model organisms aids in understanding for example metabolic disorders in human. The dissertation considers modelling of metabolism and ^{13}C-labelling for quantitative analysis of metabolic fluxes in yeast *Saccharomyces cerevisiae* that is an important biotechnological production and model organism, and in yeast *Pichia pastoris* and in fungus *Trichoderma reesei* that serve as efficient hosts for protein production.

ISBN 978-951-38-7371-4 (soft back ed.)
ISSN 1235-0621 (soft back ed.)

ISBN 978-951-38-7372-1 (URL: http://www.vtt.fi/publications/index.jsp)
ISSN 1455-0849 (URL: http://www.vtt.fi/publications/index.jsp)