THESIS ON NATURAL AND EXACT SCIENCES B171

Absolute Quantitative Multi-omics Characterization of Specific Growth Rate-dependent Metabolism of *Escherichia coli*

KASPAR VALGEPEA

TUT PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
Faculty of Science
Department of Chemistry

This thesis was accepted for the defence of the degree of Doctor of Philosophy (in Chemistry) on May 29th, 2014

Supervisor: Professor Raivo Vilu, PhD,
Department of Chemistry,
Tallinn University of Technology, Estonia

Co-supervisor: Senior Scientist Kaarel Adamberg, PhD,
Department of Food Processing,
Tallinn University of Technology, Estonia

Opponents: Professor Kazuyuki Shimizu, PhD,
Institute of Advanced Biosciences,
Keio University, Japan

Tobias Maier, PhD,
Maier Mateu Science Communication, Spain

Defence of the thesis: July 9th, 2014 in Tallinn, Estonia

Declaration:
I hereby declare that this doctoral thesis, submitted for the doctoral degree at Tallinn University of Technology, is my original investigation and achievement and has not been submitted for the defence of any other academic degree elsewhere.

_____
Kaspar Valgepea

ISSN 1406-4723
ISBN 978-9949-23-645-9 (publication)
ISBN 978-9949-23-646-6 (PDF)

LOODUS- JA TÄPPISTEADUSED B171

Absoluutselt kvantitatiivsetel oomikameetoditel põhinev kasvuerikiirusest sõltuva *Escherichia coli* metabolismi kirjeldamine

KASPAR VALGEPEA

TTÜ
KIRJASTUS

CONTENTS

ABSTRACT

Metabolism of microorganisms manifests remarkable flexibility by adaptation to different environmental conditions. This is usually accompanied by a change in the specific growth rate (μ), which integrates regulation of many properties of cell physiology *e.g.* cell morphology, proteome content, metabolic flux patterns. As the majority of these processes are regulated through coordinated changes in gene expression, a systems biology approach is needed to achieve a quantitative whole-cell level understanding of the complex control principles of metabolism. The aim of this thesis is to fully quantitatively characterize μ-dependent multi-level dynamics of steady state metabolism of the most-studied bacterium *Escherichia coli* using a systems biology approach by integrating absolute quantitative multi-omics analyses, advanced continuous cultivation and computational methods, with a special focus on acetate overflow metabolism and control of protein and metabolic flux levels.

Detailed carbon balance analysis in glucose-limited accelerostat (A-stat) cultures revealed novel carbon wasting profiles into pyrimidine pathway intermediates and together with metabolic flux analysis (MFA) suggested that despite the ~4-fold increased carbon wasting with μ rising from 0.1 to 0.5 h^{-1}, *E. coli* could maintain a constant biomass yield through a simultaneous 36% reduction of non-growth associated ATP production. Furthermore, high-resolution A-stat cultures precisely captured the start of acetate overflow at $\mu = 0.27\ h^{-1}$ and together with integrated quantitative exo-metabolome, transcriptome and proteome analyses coupled to MFA allowed to propose that acetate overflow is triggered by carbon catabolite repression-mediated down-regulation of acetyl-CoA synthetase (Acs) resulting in decreased assimilation of acetate produced by phosphotransacetylase (Pta) and disruption of the PTA-ACS node. This was experimentally proven by creating an *E. coli* Δ*pka* Δ*arcA* strain with postponed and 4-fold reduced acetate overflow through increased acetate recycling capabilities and downstream throughput in the TCA cycle. In addition, this strain is superior to previous acetate overflow-reduced *E. coli* strains since it does not accumulate any other detrimental by-product and maintains μ_{max} of wild-type.

High correlations (R ~ 0.8) between both genome-wide mRNA and protein expression changes with rising μ and concentrations were detected. Integration of absolute quantitative transcriptome and proteome data with flux rates and statistical covariance analysis revealed that *E. coli* achieved 5-time faster growth through 3.7-fold increased apparent *in vivo* catalytic rates of enzymes (k_{app}) and 2.8-fold increased translation rates by predominantly controlling protein abundances and flux rates at post-transcriptional and post-translational levels, respectively. Our analysis further showed that *E. coli* invests most of its proteome resource into expression of proteins involved in biosynthesis and energy generation while enzymes carrying high fluxes seem to be more abundant and also possess higher k_{app}.

This thesis contributes to the much-needed better quantitative description and understanding of biological systems and control principles of cell metabolism at the whole-cell level. Additionally, it advances novel modeling approaches and metabolic engineering of superior microbial cell factories.

KOKKUVÕTE

Mikroorganismide metabolismi märkimisväärset paindlikkust tõestab nende võime kohaneda erinevate keskkonnatingimustega. Reeglina kaasneb sellega muutus rakkude kasvuerikiiruses (μ), mis integreerib mitmete rakufüsioloogia mehhanismide regulatsiooni (nt. raku morfoloogia, proteoomi sisaldus, metaboolsete voogude muster). Kuna enamus neid protsesse reguleeritakse läbi koordineeritud geenide ekspressiooni muutuste, siis on vajalik rakendada süsteemide bioloogia lähenemist, et mõista komplekseid metabolismi regulatsioonimehhanisme kvantitatiivselt terve raku tasemel. Käesoleva väitekirja eesmärk on kirjeldada kvantitatiivselt enim uuritud bakteri *Escherichia coli* μ-sõltuvat *steady state* metabolismi mitmetasandilist dünaamikat kasutades süsteemide bioloogia lähenemist läbi absoluutselt kvantitatiivsete oomika-, kõrgetasemeliste läbivoolukultivatsiooni- ja arvutusmeetodite integreerimise, omades erilist fookust atsetaadi ülevoolu metabolismil ning valkude tasemete ja voogude suuruste kontrollmehhanismidel.

Detailne süsinikubilansi analüüs glükoos-limiteeritud akselerostaat (A-staat) kultuurides võimaldas tuvastada uudsed pürimidiini raja vaheühendite süsiniku raiskamisprofiilid ning koos metaboolsete voogude analüüsiga (MFA) viitas sellele, et hoolimata *ca* nelja-kordsest süsiniku raiskamise kasvust μ tõusmisel 0.1-lt 0.5-le h^{-1}, suutis *E. coli* säilitada konstantse biomassi saagise tänu samaaegsele 36%-lisele kasvuga mitte-seotud ATP tootmise vähendamisele. Veel enam, kõrge resolutsiooniga A-staat kultuuride kasutamine võimaldas täpselt tuvastada atsetaadi ülevoolu alguse μ = 0.27 h^{-1} juures ning koos integreeritud kvantitatiivsete rakuvälise metaboloomi, transkriptoomi ning proteoomi analüüside ja MFA-ga võimaldas püstitada hüpoteesi, et atsetaadi ülevoolu käivitab kataboliitse repressiooni poolt põhjustatud atsetüül-CoA süntetaasi (Acs) allaregulatsioon, millega kaasneb fosfotransatsetülaasi (Pta) poolt toodetud atsetaadi tarbimisvõime langus ning PTA-ACS tsükli katkemine. Selle hüpoteesi eksperimentaalseks tõestamiseks loodi *E. coli* Δ*pka* Δ*arcA* tüvi, mille puhul neli korda metsiktüvest väiksem atsetaadi ülevool algab ka oluliselt kõrgemal glükoosi tarbimiskiirusel tänu kõrgemale atsetaadi tagasitarbimis- ning edasisele metaboliseerimisvõimele TCA tsüklis. Lisaks on tüvi parem eelnevatest vähendatud atsetaadi ülevooluga *E. coli* tüvedest, kuna ei tooda ühtegi teist kahjulikku kõrvalprodukti ja omab metsiktüvedega võrdset μ_{max}-i.

Väitekirjas detekteeriti kõrged korrelatsioonid (R ~ 0.8) nii üle-genoomsete mRNA ja valgu ekspressiooni μ-sõltuvate muutuste kui ka kontsentratsioonide vahel. Absoluutselt kvantitatiivsete transkriptoomi ja proteoomi andmete integreerimine voogude kiiruste ning statistilise kovariatsioonianalüüsiga tuvastas, et *E. coli* saavutas 5-korda kiirema kasvu läbi 3.7-korda tõusnud näivate *in vivo* ensüümide katalüütiliste (k_{app}) ja 2.8-korda kasvanud translatsiooni kiiruste, kontrollides peamiselt valkude koguseid ning voogude kiirusi vastavalt post-transkriptsiooni ja post-translatsiooni tasemetel. Lisaks täheldati, et *E. coli* investeerib enamuse proteoomi ressursist biosünteesi ja energia tootmisega seotud valkude ekspresseerimisele ning et suuri vooge katalüüsivaid ensüüme leidub rakus nii suuremal arvul, kui ka neil on kõrgemad k_{app} väärtused.

Käesolevas väitekirjas kogutud andmed võimaldavad oluliselt paremini kui varem kirjeldada kvantitatiivselt metabolismi regulatsioonimehhanisme terve raku tasemel. Täiendavalt edendab väitekiri uudseid modelleerimismeetodeid ning paremate mikroobsete tootjarakkude konstrueerimist.

LIST OF PUBLICATIONS

This thesis is based on the following publications referred to in the text by their Roman numbers and reproduced in the appendices with permission from the publishers.

I **Valgepea K**, Adamberg K, Nahku R, Lahtvee P-J, Arike L & Vilu R (2010) Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase. *BMC Syst. Biol.* 4, 166.

II **Valgepea K**, Adamberg K & Vilu R (2011) Decrease of energy spilling in *Escherichia coli* continuous cultures with rising specific growth rate and carbon wasting. *BMC Syst. Biol.* 5, 106.

III Peebo K, **Valgepea K**, Nahku R, Riis G, Õun M, Adamberg K & Vilu R (2014) Coordinated activation of PTA-ACS and TCA cycles strongly reduces overflow metabolism of acetate in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 98, 5131–5143.

IV **Valgepea K**, Adamberg K, Seiman A & Vilu R (2013) *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins. *Mol. Biosyst.* 9, 2344–2358.

Author's Contribution to the Publications

I The author designed and coordinated the project, designed and performed the experiments, performed exo-metabolome analysis and aided in other –omics analyses, analyzed the data, and wrote the manuscript.

II The author designed and coordinated the project, designed and performed the experiments, performed all analyses, analyzed the data, and wrote the manuscript.

III The author designed and coordinated the project, participated in designing and performing the experiments, analyzing the data and in writing the manuscript.

IV The author designed and coordinated the project, performed calculations of absolute transcriptome and proteome data, participated in developing data analysis methods, data integration and analysis, and wrote the manuscript.

Additional publications

V Adamberg K, Lahtvee P-J, **Valgepea K**, Abner K & Vilu R (2009) Quasi steady state growth of *Lactococcus lactis* in glucose-limited acceleration stat (A-stat) cultures. *Antonie Van Leeuwenhoek* 95, 219–226.

VI Lahtvee P-J, **Valgepea K**, Nahku R, Adamberg K, Abner K & Vilu R (2009) Steady state growth space study of *Lactococcus lactis* in D-stat cultures. *Antonie Van Leeuwenhoek* 96, 487–496.

VII Nahku R, **Valgepea K**, Lahtvee P-J, Erm S, Abner K, Adamberg K & Vilu R (2010) Specific growth rate dependent transcriptome profiling of *Escherichia coli* K12 MG1655 in accelerostat cultures. *J. Biotechnol.* 145, 60–65.

VIII Nahku R, Peebo K, **Valgepea K**, Barrick JE, Adamberg K & Vilu R (2011) Stock culture heterogeneity rather than new mutational variation complicates short-term cell physiology studies of *Escherichia coli* K-12 MG1655 in continuous culture. *Microbiology* 157, 2604–2610.

IX Arike L, **Valgepea K**, Peil L, Nahku R, Adamberg K & Vilu R (2012) Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *J. Proteomics* 75, 5437–5448.

X Pey J, **Valgepea K**, Rubio A, Beasley JE & Planes FJ (2013) Integrating gene and protein expression data with genome-scale metabolic networks to infer functional pathways. *BMC Syst. Biol.* 7, 134.

LIST OF CONFERENCE PRESENTATIONS

I **Valgepea K**, Adamberg K, Seiman A & Vilu R. Absolute quantitative triple-omics analyses reveals that *Escherichia coli* achieves faster growth by increasing catalytic activities of enzymes and translation rate.
Oral presentation at *Microbial Stress: from Molecules to Systems, May 2012, Belgirate, Italy*

II **Valgepea K**, Adamberg K, Nahku R, Lahtvee P-J, Arike L & Vilu R. Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase.
Oral and Poster presentations at *Annual Symposium of the Danish Microbiological Society, November 2010, Copenhagen, Denmark*

III **Valgepea K**, Adamberg K, Nahku R, Lahtvee P-J, Arike L & Vilu R. Growth rate dependent acetate metabolism characterization in *Escherichia coli* continuous cultures.
Oral presentation at *Recent Advances in Fermentation Technology VIII, November 2009, San Diego, USA*

IV **Valgepea K**, Adamberg K, Seiman A & Vilu R. *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins.
Poster presentation at *New Approaches and Concepts in Microbiology, October 2013, Heidelberg, Germany*

V **Valgepea K**, Adamberg K & Vilu R. Systems biology approach reveals decrease of energy spilling in *Escherichia coli* continuous cultures with rising specific growth rate and carbon wasting.
Poster presentation at *Society of Industrial Microbiology 2011 Annual Meeting and Exhibition, July 2011, New Orleans, USA*

VI **Valgepea K**, Adamberg K & Vilu R. Novel carbon wasting profiles and metabolic flux dynamics in *Escherichia coli* are dependent on specific growth rate.
Poster presentation at *Joint FEBS/SystemsX Advanced Lecture Course on Systems Biology, February 2011, Innsbruck, Austria*

VII **Valgepea K**, Adamberg K, Nahku R, Lahtvee P-J, Arike L & Vilu R. Overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase.
Poster presentation at *11th International Conference on Systems Biology, October 2010, Edinburgh, Scotland*

INTRODUCTION

Escherichia coli is the most-studied bacterium and also widely exploited by the biotechnology industry as a microbial cell factory. Hence, there is great interest from both the industry and academia to gain even more knowledge about the metabolism of this gram-negative model-organism. Whereas a lot about *E. coli* metabolism is known, there exist gaps or lack of knowledge altogether regarding some fundamental principles of metabolic regulation.

For example, although the specific growth rate (μ) is known to be central in integrating regulation of many properties of cell physiology *e.g.* cell morphology [1–3], gene expression [4–8], metabolic flux patterns [4,7–10], simultaneous regulation of cellular metabolome, transcriptome, proteome and fluxome levels and their patterns accompanying the change in μ is not thoroughly characterized and understood. More specifically, for example, after decades of studies the mechanism behind the regulation of acetate overflow metabolism has not been unequivocally elucidated. Furthermore, regulation levels of the gene expression cascade (*e.g.* transcriptional, post-translational) controlling protein abundances and metabolic fluxes enabling the cells to achieve faster growth are not known. To gain new insights into these phenomena and μ-dependent metabolism in general, the main aim of this thesis is to comprehensively and fully quantitatively characterize μ-dependent multi-level dynamics of steady state metabolism of the most-studied bacterium *E. coli* using a systems biology approach by coupling absolute quantitative multi-omics analyses with advanced continuous cultivation and computational methods. It is important to point out that it is critical that this biological information is acquired for cells grown in strictly defined physiological states *i.e.* continuous cultures, representing steady state physiology [11,12].

This thesis advances the much-needed better quantitative description and understanding of biological systems and control principles of cell metabolism at the whole-cell level [13,14] in many aspects. For instance, we report novel μ-dependent carbon wasting profiles; propose a new mechanism for acetate overflow metabolism and experimentally prove this by creating strains with reduced carbon loss into acetate; determine that *E. coli* achieves faster growth by increasing catalytic and translation rates of proteins.

ABBREVIATIONS

A-stat	accelerostat
Acetyl-P	acetyl phosphate
AckA	acetate kinase
Acs	acetyl-CoA synthetase
ArcA	dual transcriptional regulator for anoxic redox control
ATP	adenosine triphosphate
cAMP	cyclic AMP
CBASP	carbamoyl-aspartate
CCM	central carbon metabolism
CCR	carbon catabolite repression
Crp	cAMP receptor protein
D	dilution rate
D-stat	dilution rate stat
DHO	dihydroorotate
DNA	deoxyribonucleic acid
E. coli	*Escherichia coli*
gDCW	gram dry cellular weight
GS	glyoxylate shunt
k_{app}	apparent *in vivo* catalytic rate of enzyme
k_{cat}	the maximum number of reactions catalyzed per enzyme
MFA	metabolic flux analysis
mRNA	messenger RNA
NAA	acetyl-aspartate
NADH	reduced nicotineamide adenine dinucleotide
NADPH	reduced nicotineamide adenine dinucleotide phosphate
nATP	protein synthesis cost
PDH	pyruvate dehydrogenase
PEP	phosphoenolpyruvate
Pka	protein lysine acetyltransferase
pm	protein-per-mRNA ratio
Pta	phosphotransacetylase
PPP	pentose phosphate pathway
PoxB	pyruvate oxidase
R	Pearson correlation coefficient
RNA	ribonucleic acid
TCA cycle	tricarboxylic acid cycle
TR	transcriptional
TL	translational
Y_{XS}	biomass yield
μ	specific growth rate

1. LITERATURE REVIEW

1.1. Importance of *Escherichia coli*

Escherichia coli is a gram-negative, rod-shaped and non-sporulating bacterium commonly found in the lower-intestines of warm-blooded animals. It is a facultative aerobe meaning it can easily switch between fully respiratory, fermentative and respiro-fermentative growth. *E. coli* was first discovered in 1885 by a German bacteriologist Theodor Escherich. Notably, it has been the most-studied prokaryotic model organism both in microbiology and biotechnology. It could also be considered the most important free-living organism since a significant part of the knowledge regarding fundamentally conserved biological processes in living organisms at the molecular level derive from laboratory studies of *E. coli*. Furthermore, this bacterium also played a crucial role in the start of biotechnology by serving as the host for the first recombinant DNA technology invented by Genentech, Inc. Since then, *E. coli* has been vastly and successfully exploited in the biotechnology industry for the production of recombinant proteins, low molecular weight compounds, biofuels etc., mainly because of low manufacturing and end-product purification costs and its ability to easily reach high cell densities [15,16]. Due to the latter and the notion that new information regarding *E. coli* can be more successfully interpreted because of the vast amount of knowledge which already exists, studying *E. coli* still has high relevance for further advancement of our fundamental understanding of living organisms and the needed innovation in the biotechnology industry.

Metabolism of *E. coli* is highly flexible which facilitates successful adaptation to a wide range of environmental conditions. For example, *E. coli* can rapidly switch from growing on amino acid substrates to synthesizing all the 20 amino acids necessary for biomass proliferation if substrates are suddenly depleted. Furthermore, it can also grow aerobically or anaerobically, survive osmotic or acidic stress etc. Realization of such flexibility requires sophisticated and tightly controlled regulation mechanisms, about which we still lack a complete understanding. Hence, any additional knowledge regarding the regulation principles of cell metabolism are highly valuable for a more comprehensive understanding of biological systems overall [13,14] and novel and more successful biotechnology and synthetic biology efforts [17]. As adaptation to different environmental conditions is usually accompanied by a change in the specific growth rate (μ), which integrates regulation of many properties of cell physiology *e.g.* cell morphology [1–3], gene expression [4–8], metabolic flux patterns [4,7–10], studying the metabolic responses of *E. coli* to perturbations in μ at the most important regulatory layers of transcriptome, proteome and fluxome can lead to elucidation of the complex control principles of metabolism at whole-cell level.

1.2. Specific growth rate-dependent steady state metabolism of *Escherichia coli*

1.2.1. Importance of steady state cultures

The most widely-used method for studying cell growth is batch cultivation. Indeed, batch cultures are very effective for high-throughput screening of μ and maximum cell yield in microtiter plates and shake flasks. However, batch cultures result in non-steady state growth even during the so-called 'balanced growth phase' (when substrate consumption is maximal during exponential growth) due to the constantly changing biomass and metabolic by-product concentrations. Moreover, trying to study μ-dependent metabolism in batch cultures using different media does not allow one to draw unequivocal conclusions regarding μ-dependent metabolism since manipulating μ by varying substrates also leads to substrate specific re-organization of metabolism. It has been concluded that batch cultures result in complex data patterns reflecting uncontrolled changes of growth conditions, which are often difficult or even impossible to interpret [11,12], making them unsuitable for studying μ-dependent metabolism.

To overcome the issues with batch cultures, the continuous cultivation method chemostat was introduced in 1950 to study μ-dependent metabolism in steady state using strictly defined and controlled growth conditions [18,19]. As studying metabolism over a wide range of μ with high-resolution using chemostat cultures is very time and resource consuming, the continuous cultivation method accelerostat (A-stat) was later developed [3]. A-stat enables to collect vast amount of data in short time to monitor the dynamics of metabolism with very high resolution and precisely detect metabolic switch-points (*e.g.* start of acetate overflow), which could be left unnoticed using chemostats. Importantly, quasi steady state data of A-stat describe steady state physiology equally to chemostats if experiments are conducted properly [20–26]. In conclusion, it is important that μ-dependent metabolism is studied in strictly defined physiological states *i.e.* continuous cultures, representing steady state physiology [11,12]. Thus all the following literature review under this sub-heading considers continuous cultures, and more specifically glucose-limited cultures.

1.2.2. Carbon flows

Carbon is the central element for every organism and understanding its metabolism also in *E. coli* is instrumental for understanding its physiology. Majority of the consumed carbon by aerobically growing *E. coli* is used for biomass formation and CO_2 synthesis [4–7,9,10,27–32], which is mostly formed as an accompanying by-product for generating energy and reducing equivalents. Additionally, a notable amount of the consumed carbon is lost to several by-products excreted by the cells into the growth environment. The main by-product in wild-type *E. coli* aerobic cultivations is acetic acid [4–10,31–34]. In addition, accumulation of other compounds such as lactate, formate, pyruvate, ethanol etc. has been observed [4,10,35]. Although

excretion of many other compounds besides the above-mentioned 'well-known' ones has been detected [4,36–38] *e.g.* pyrimidine pathway intermediates, these carbon wasting outflows are generally not taken into account in carbon balance or metabolic flux analysis (MFA) of *E. coli*, possibly leading to questionable conclusions. For instance, Taymaz-Nickerel and colleagues accounted a substantial amount of 'missing carbon' in the carbon balance (7–13%) of *E. coli* continuous cultures to cells lysis which has not been observed before in the literature [39].

As the above-mentioned carbon flows and their relative fractions from the consumed carbon vary with growth conditions, it is very important to quantitatively determine the dependence of carbon flows on μ. Glucose-limited continuous culture experiments with various *E. coli* wild-type strains have generally shown a constant or slightly increased carbon flow to biomass and decreased fraction allocated for CO_2 synthesis with rising μ [4,5,7,9,10,27,28,31,40]. Regarding the major by-product acetate, it is quite well documented that acetate overflow is μ-dependent with no excretion at slow and high excretion at fast growth [4–10,31–34]. The μ-dependent excretion patterns of other by-products are much less clear or missing altogether. Therefore, to obtain a more accurate description of μ-dependent carbon flows and their regulation in the metabolic network, A-stat experiments with high-resolution of μ together with detailed carbon balance analysis should be carried out.

1.2.3. Metabolic fluxes and energy metabolism

Knowing carbon uptake and excretion routes gives a general understanding about carbon catabolism but a more sophisticated method of flux analysis (fluxomics) is necessary to quantify metabolic fluxes and their distribution within the metabolic network. However, metabolic fluxes cannot be measured directly *in vivo*, but can be estimated either with computational methods (*e.g.* MFA, flux balance analysis) or experimental flux measurement techniques with stable isotope tracers (*e.g.* ^{13}C-MFA) [41,42]. These methods require a thorough knowledge of the metabolic network, its stoichiometry and biomass composition. However, balancing the energy-metabolites ATP, NADH and NADPH is difficult since *E. coli* possesses several transhydrogenases able to interconvert NADH and NADPH [9,43], there exist ATP dissipating futile cycles [28,44–48] and the exact efficiency of NADH to ATP conversion (P/O) in the respiratory chain (RC) is not known [49,50]. Also the biomass composition is dependent on growth conditions [27,39,51–53] and using the one corresponding to the experimental state is important since it influences flux calculations as shown by sensitivity analysis [52]. These challenges could be met by using additional constraints based on ^{13}C-MFA [4,7,9] and experimentally determined biomass composition [52], making fluxomics methods acceptable for predicting metabolic flux patterns and studying energy metabolism also in *E. coli* [41,42].

It is difficult to conclude about the exact splitting ratio of initial glucose catabolism between glycolysis and pentose phosphate pathway (PPP) in *E. coli* due to the uncertainty about transhydrogenase activities *in vivo* [9,43]. Still, it seems that roughly a 70 to 30% splitting between glycolysis and PPP occurs in *E. coli* glucose-limited

continuous cultures [4,7–10,28,54] and this ratio changes in favor for PPP with rising μ, presumably to meet the increasing NADPH demands for higher RNA synthesis with faster growth [55–58]. In parallel, tricarboxylic acid (TCA) cycle fluxes decrease with the concomitant start and increasing acetate overflow with rising μ. Strong μ-dependent activity of the glyoxylate shunt (GS) has also been determined with high activity at low μ and no flux above $\mu = 0.4\ h^{-1}$ [4,7,9]. Simultaneously, the anaplerotic PEP carboxylase flux increases with rising μ.

Quantification of metabolic flux patterns is essential for understanding the regulation of energy metabolism. Although balancing energy-metabolites accurately to reflect the true situation *in vivo* is difficult to achieve, it is generally accepted that around half of the produced energy is spent for functions not directly related to growth [31,44,45,59,60]. Probably the most energy-demanding processes of the latter are turnover of macromolecules, re-establishment of ion gradients and futile cycles [44,45,58,60,61]. This non-growth related energy production is also termed maintenance energy and usually further divided into non-growth and growth associated parts. However, these terms should be used with great caution due to the confusion regarding their exact biological nature [44,45,60]. The non-growth associated maintenance cost is assumed to be constant and generally estimated by extrapolating the specific glucose consumption rates (q_{glc}) measured in chemostat cultures to $\mu = 0\ h^{-1}$ [9,40,44,45,60]. However, since the total non-growth related energy costs are known to vary among growth conditions [44,45,60,62,63], constant values should be avoided and regulation of μ-dependent energy metabolism should be inferred from the net difference between energy production and consumption, assuming that the P/O is insensitive to μ [39].

1.2.4. Functional-genomic responses

Adaptation of microorganisms to different environmental conditions is generally accompanied by a change in μ, which integrates regulation of many properties of cell physiology *e.g.* cell morphology [1–3], metabolic flux patterns [4,7–10]. As the majority of these mechanisms are regulated through changes in gene expression, it is very useful to study the functional-genomic responses of *E. coli* to a perturbation also in μ to gain further insights into the complex control principles of metabolism. The most available methods for this are transcriptomics and proteomics.

One could expect activation of gene expression in pathways responsible for the production of precursors for biosynthesis and energy generation to realize the necessary significantly higher flux throughput of these pathways for faster growth. Although the studies of μ-dependent genome-wide transcriptome analysis in the literature differ by *E. coli* wild-type strains, media and other conditions [4–7], the results in general do not exactly support the latter expectation: mRNA levels in glycolysis are rather constant or slightly increase with rising μ and in the TCA cycle increase but drop off at faster growth. PPP transcript levels, however, uniformly increase with faster growth. These observations at mRNA level are mostly consistent with the only μ-dependent proteome data set [4]. Data regarding mRNA expression of

the components of the RC differ between the studies as constant, decreasing and increasing patterns have been observed with rising μ. Interestingly, however, all the studies detected a strong increase of *ndh*, an NADH dehydrogenase not generating proton motive force [64], with faster growth, possibly linked to balancing the redox ratio which has been proposed to be an issue in faster growing *E. coli* [5].

Another common observation among the latter studies is the activation of carbon catabolite repression (CCR) [65,66] with rising μ, evidenced by down-regulation of the genes associated with alternative (to glucose) substrate transport and catabolism [4–7]. This could be expected as the rising residual glucose concentration in glucose-limited continuous cultures with faster growth could trigger CCR, which realizes the preference of glucose over alternative carbon sources in the presence of glucose through the repression of genes necessary for the use of alternative carbon sources [65,66]. Closely coupled with this regulation is the strong down-regulation of GS and gluconeogenesis at both mRNA and protein levels [4–7]. Consistent with the results of flux analysis described above is the up-regulation of the anaplerotic PEP carboxylase (*ppc* and Ppc) with rising μ in all the studies, presumably activated to replenish the decreasing pools of TCA cycle intermediates oxaloacetate and α-ketoglutarate due to increasing demands of these precursor molecules for amino acid and nucleotide biosynthesis with faster growth [58].

Recent reviews [67–70] have concluded that time-course analyses of bacteria and yeast reveal large differences between mRNA and protein expression changes in perturbed non-steady state batch cultures, indicating considerable post-transcriptional (post-TR) regulation. Notably, the only *E. coli* steady state data set with simultaneous transcript and protein profiling, though for only 56 genes, shows correlations up to R = 0.4 (Pearson correlation coefficient) between mRNA and protein changes with rising μ [4]. The different results between non-steady state and steady state cultures highlights the need for a more global analysis at both levels in steady state conditions to investigate if the state of the culture could be an important factor in mRNA-protein correlation analysis and for understanding the relevance of post-TR regulation.

1.3. Acetate overflow metabolism of *Escherichia coli*

As stated above, the main by-product for wild-type *E. coli* aerobic cultivations is acetic acid [4–10,31–34]. Acetic acid exerts its toxicity by uncoupling the transmembrane pH gradient and acidifying the cytoplasm [71–73]. In addition to being detrimental for recombinant protein synthesis, acetate interferes with methionine biosynthesis and its accumulation diverts valuable carbon from biomass formation and inhibits growth even at as low concentrations as 0.5 g/L [15,74–78].

Acetate excretion into the growth environment is also termed acetate overflow and it is known to be μ-dependent with no excretion at slow and high excretion at fast growth [4–10,31–34]. Acetate overflow metabolism has been studied widely over the years and it is generally believed to be caused by an imbalance between substrate uptake and anabolic/catabolic throughput of downstream pathways [79,80]. Several explanations propose limitations in respiratory capacity [31,32,79], TCA cycle

throughput [81,82], energy generation [10,72], activity of the GS [33,83], necessity for coenzyme A replenishment [84] or membrane space [85]. Numerous attempts at process and genetic levels have been tried to diminish acetate overflow (reviewed in [86]). Notably, deletion of the main synthesis pathways phosphotransacetylase-acetate kinase (PTA-ACKA) and pyruvate oxidase (POXB) [72] reduce acetate overflow but result in several detrimental trade-offs: reduction of μ and biomass yield (Y_{XS}), and substantially elevated carbon loss into lactate and formate [29,76,87–90]. This shows that acetate overflow cannot be simply avoided by removing the main synthesis pathways. It is actually not surprising that acetate overflow cannot be totally abolished with this approach as down-regulation of the main acetate synthesis pathway genes is observed in chemostats before the start of acetate overflow [4–6,33]. In conclusion, none of the general theories or process/genetic efforts has been able to unequivocally explain the mechanism and regulation behind acetate overflow.

Interestingly, a theory for acetate overflow in *Saccharomyces cerevisiae* proposes that acetate accumulation is the result of insufficient acetyl-CoA synthetase (Acs) activity for the complete functioning of the pyruvate dehydrogenase bypass because of glucose repression of Acs at high μ [91]. It seems that this theory could also be relevant in *E. coli* since acetate overflow starts at a lower μ in an *E. coli acs* knockout strain compared to wild-type [33]. Although, no clear conclusions can be drawn from literature data with single over-expression of *acs* [33,92], its over-expression together with deletion of GS repressors *iclR* and *fadR* reduces acetate excretion in batch cultures [33]. Furthermore, similar to *S. cerevisiae*, CCR [65,66] is responsible for repression of Acs activity in *E. coli* with rising μ [4–6] since acetate is an alternative carbon source to glucose. The theory that CCR-mediated repression of *acs* could trigger acetate overflow is supported by the observation that an *E. coli* knockout strain of *crp* (a central player in CCR of *E. coli* through activating expression of catabolic genes such as *acs* in complex with cyclic AMP (cAMP) [65,66,93]) accumulates acetate also at low μ where wild-type does not [7,94]. Moreover, *E. coli* CRP* mutants that do not require Crp binding to cAMP to activate the expression of catabolic genes showed ~4-time higher *acs* expression and secreted substantially less acetate in xylitol producing batch fermentations [95]. Of course, other mechanisms can be involved in Acs repression, as, for example, an *E. coli cra* knockout strain shows increased acetate production rates [96]. Still, it can be concluded that after decades of studies, the mechanism and regulation behind acetate overflow metabolism of *E. coli* still remain unclear.

1.4. Absolute quantitative –omics analyses

With the recent rapid advances in high-throughput –omics analyses such as transcriptomics and proteomics [68], analysis of cell metabolism at absolute quantitative levels—molecules per cell or intracellular concentrations—has become a reality. The importance of absolute quantification can be illustrated by the following example: if a cell increases expression of protein X 10-fold as a response to an environmental perturbation, it does not necessarily mean that the energetic and

expression load of protein X on metabolism is more significant than that of protein Y which is up-regulated only 2-fold. This is because the real metabolic load of protein expression is determined also by the protein amount and that metabolic load would be much higher for protein Y if its abundance is 10-fold higher than that of protein X. Furthermore, since translation is believed to be the rate-limiting process for faster growth, at least in *E. coli* [97], it is highly relevant to analyze the allocation principles of proteome resources for understanding μ-dependent metabolic regulation [97,98].

Simultaneous absolute quantification of mRNA and protein levels allows to generate protein-per-mRNA ratios (pm) which estimate translation efficiencies and changes in pm values inform about the level of gene expression regulation (*e.g.* TR, post-TR) either through protein translation or degradation [69]. Interestingly, abundant proteins seem to be translated more efficiently than those of low abundance [99,100]. Translation efficiency can be maximized through codon bias *i.e.* non-random occurrence of codons for coding amino acids [67,101], which shows genome-wide correlation with protein expression levels [102–107]. In addition, coupling absolute proteome quantification with protein turnover measurements [99,108] opens a new dimension for metabolic engineering of superior cell factories as different routes/pathways to the target product can be evaluated in terms of their translational load and effects of protein re-synthesis costs on energy homeostasis.

Genome-wide absolute quantification of transcriptome and proteome has been carried out for several microorganisms in batch cultures: proteome in *E. coli* [102,104,105,109] and simultaneous transcriptome and proteome in *E. coli* [110], *Bacillus subtilis* [111], *Mycoplasma pneumoniae* [108] and yeast [100]. However, batch cultures do not reflect steady state physiology as highlighted above. The only absolute quantitative studies conducted in steady state continuous cultures of *E. coli*, to the best of our knowledge, are those of genome-wide transcriptome [40] and simultaneous quantification of concentrations of 56 transcripts and proteins [4] at various μ. The latter study shows correlations of $R \sim 0.8$ between mRNA and protein concentrations. Despite these studies, we still miss μ-dependent genome-wide absolute quantitative mRNA and proteome data for *E. coli*. This data is valuable for a better quantitative understanding of cellular processes at whole-cell level [13,14] and for novel whole-cell modeling approaches where cell metabolism is simulated also as a function of μ [112–115].

1.5. Systems biology study of metabolism

Most biological studies look at a specific aspect or layer of metabolism. However, to achieve a more systems level (whole-cell) understanding of physiological processes and metabolic regulation, a systems biology approach of integrating transcriptome, proteome and fluxome data coupled to models of different levels of detail is needed [14,68,116,117]. An effort to capture a whole-cell snapshot of both the components and the interactions between them is a tremendous challenge, but as understanding of both the regulation levels of protein and flux levels and other molecular relationships enabling the cells to modify μ is of instrumental importance towards a more complete

description of the control principles of cell metabolism [13,14], more accurate modeling [112–115] and successful biotechnology and synthetic biology efforts [17], genome-wide absolute quantitative multi-omics analysis in one study is highly needed.

Although several research groups have recently approached this challenge in various microorganisms [108,111,118–121], each study either lacks one necessary layer or the cells were not cultivated in steady state cultures. Similarly, many μ-dependent cellular global relationships have been determined in *E. coli* non-steady state batch cultures [55] and at single layers of transcriptome [5–8,40] or fluxome [7–9,28] in continuous cultures. There is one landmark study of *E. coli*, however, which captures all the –omics levels at various μ in steady state chemostat cultures [4]. Still, this seminal work lacks a proteome-wide view as μ-dependent patterns were determined for only 56 proteins. Thus there is still a need for a study which captures μ-dependent global responses of absolute quantitative transcriptomes and proteomes of *E. coli* together with flux analysis, potentially leading to elucidation of the complex control principles of metabolism at whole-cell level.

One very interesting question what such integrated systems biology studies of metabolism could address is: at which regulation levels of the gene expression cascade are protein and flux levels controlled (*e.g.* TR, post-TR, translational (TL)) in different metabolic pathways for achieving higher flux throughput and faster growth? Determination of metabolic flux control levels is especially important for understanding regulation of metabolism since they represent the integrated response of all levels of cellular regulation [41]. Gene expression regulation of enzyme and following metabolic flux levels is not straightforward: enzyme abundance can be regulated through different regulation of mRNA and protein degradation, protein translation, post-TL modification, possible functional requirement for protein binding [67–70]; flux throughput by post-TL regulation of the catalyzing enzyme (either through chemical modification or allosteric regulation) or change of its catalytic rate through hyperbolic change of enzyme kinetics solely due to substrate concentration changes [122].

Although we are missing information regarding μ-dependent control levels of protein and fluxes in *E. coli*, recent data uniformly reveal predominant post-TR regulation of protein abundances in bacteria with very different growth characteristics: *M. pneumoniae* [108], *B. subtilis* [111] and *Lactococcus lactis* [123]. Hierarchical regulation analysis has mainly been used as the methodology for determining to which extent a change in a particular flux is regulated at gene expression or post-TL levels [124]. Most of the studies implementing this approach have studied yeast and concluded that fluxes through glycolytic and fermentative pathways are mainly regulated at the post-TR level (reviewed in [125]). Notably, the only μ-dependent data set shows predominant post-TL control of fluxes in *L. lactis* continuous cultures [123]. Indeed, it has been concluded that changes in metabolic flux patterns are not a straightforward consequence of TR regulation of enzyme levels [116,122,126,127]. In conclusion, a comprehensive systems biology study is needed to unequivocally determine the regulation levels of protein and flux levels through which *E. coli* achieves higher flux throughput and faster growth.

Important parameters for whole-cell modeling and kinetic analysis of metabolism are the catalytic rates of enzymes (k_{cat}). However, a notable problem with the currently used k_{cat} values is that they are derived from *in vitro* enzyme assays which might not reflect the situation *in vivo* [128], even if specially developed *in vivo*-like medium is used in the assay [129]. This is where absolute quantitative systems biology can help: apparent *in vivo* catalytic rates for enzymes (k_{app}) can be estimated by directly measuring both the enzyme concentration and the flux through the enzyme by coupling proteome and fluxome analyses. This approach has been applied to estimate k_{app} values in *E. coli* [109] and *M. pneumoniae* [130] batch cultures and μ-dependent values in *L. lactis* continuous cultures [123].

In addition to the multi-omics experimental part, computational modeling is equally important for a systems biology study since large-scale data sets are very challenging to analyze without special tools. The above-mentioned rapid advances in high-throughput genome-wide –omics analyses further drives the innovation of modeling approaches by providing new data at unprecedented scale. Integration of multi-omics data with genome-scale metabolic models has recently produced totally novel whole-cell modeling frameworks [112,113,115]. These modeling approaches already have [131] and will further accelerate biological discovery while also contributing to metabolic engineering of cell factories with totally novel functions [132–135].

2. AIMS OF THE THESIS

The main aim of this thesis is to comprehensively and fully quantitatively characterize specific growth rate (μ)-dependent multi-level dynamics of steady state metabolism of the most-studied bacterium *Escherichia coli* using a systems biology approach by coupling absolute quantitative multi-omics analyses with advanced continuous cultivation and computational methods.

More specific aims are as follows:

I High-resolution description of *E. coli* μ-dependent carbon balance, metabolic flux patterns and regulation of energy metabolism in steady state.

II Gain new insights into the regulation of acetate overflow metabolism in *E. coli* and utilize this knowledge to engineer strains with reduced carbon wasting into acetate.

III For the first time, determine μ-dependent genome-wide mRNA and protein abundances and concentrations, and their dynamics with rising μ in *E. coli*.

IV Determine at which regulation levels of the gene expression cascade is control of protein abundances and metabolic fluxes realized in *E. coli* for achieving higher flux throughput and faster growth.

V Estimate apparent *in vivo* catalytic rates for *E. coli* enzymes by coupling absolute quantitative proteomics and flux analysis.

3. MATERIALS AND METHODS

Detailed descriptions of materials and methods are available in the publications forming this thesis. The following brief sections are provided to make the thesis more accessible to the reader.

3.1. Bacterial strains

Escherichia coli K-12 strains were used in cultivation experiments as follows:

Publication I, II and IV - *E. coli* K-12 MG1655 (λ- F- *rph*-1 *Fnr*+; Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), DSM No.18039)

Publication III - *E. coli* K-12 BW25113 (*lacI*^{q} *rrnB3* Δ*lacZ4787 hsdR514* Δ*(araBAD)567* Δ*(rhaBAD)568 rph-1*) and its single-gene knockout strains—Δ*acs*, Δ*pka*, Δ*cobB*, Δ*arcA*—originate from the Keio collection [136]. The single *acs* and *acs* over-expression in Δ*pka* background (Δ*pka acs*OE), and Δ*pka* Δ*arcA* double-knockout strains were constructed in the BW25113 background as described in Publication III.

3.2. Growth medium

In all accelerostat (A-stat) experiments, cells were grown in defined minimal medium supplemented with 4.5 g/L glucose with the following composition (g L^{-1}): $FeSO_4 \times 7H_2O$ 0.005, $MgSO_4 \times 7H_2O$ 0.5, $MnSO_4 \times 5H_2O$ 0.002, $CaCl_2 \times 2H_2O$ 0.005, $ZnSO_4 \times 7H_2O$ 0.002, $CoSO_4 \times 7H_2O$ 0.0006, $CuSO_4 \times 5H_2O$ 0.0005, $(NH_4)_6Mo_7O_{24} \times 4H_2O$ 0.0026 were dissolved in 50 ml 5M HCl; N-source – NH_4Cl 3.5 and buffer – K_2HPO_4 2 were autoclaved separately and mixed together afterwards. In addition, 100 or 200 µl/L of foam removing agent Antifoam C (Sigma Aldrich, St. Louis, MO, USA) was used.

The latter was also used as the main cultivation medium in dilution rate stat (D-stat) experiments. The same medium additionally supplemented with acetic acid (final concentration 5 mM) was used in one two-substrate A-stat experiment and as the second medium in D-stats to study acetate consumption capability.

3.3. Continuous cultivation conditions

Glucose-limited continuous cultivations were carried out at 37 °C, pH 7 and aerobic conditions ensured by agitation of 800 rpm and air flow rate of 150 mL/min. A-stat [3] cultures were stabilized in chemostat to achieve steady state either at dilution rate (D) 0.1 or 0.2 h^{-1} after which the A-stat phase was started with a continuous increase of D using acceleration 0.01 h^{-2}. The control algorithm for A-stat was: $D = D_0 + a_D \times t$, where D_0 is the initial D of chemostat (h^{-1}), a_D is the acceleration of D in the A-stat phase (h^{-2}) and t is the time from the start of A-stat (h).

E. coli K-12 wild-type strains were cultivated in four and three independent A-stat experiments in Publications III and I, respectively. *E. coli* K-12 MG1655 was also

cultivated in one two-substrate A-stat in Publication I. *E. coli* K-12 BW25113 mutant strains Δ*arcA* and Δ*pka* Δ*arcA* were cultivated in two while other BW25113 mutants in one independent A-stat experiment in Publication III. In D-stats, *E. coli* K-12 MG1655 was cultivated at D = 0.1; 0.24; 0.30; 0.45 and 0.51 h^{-1} in four and two independent experiments, respectively. D-stats at D = 0.19 and 0.40 h^{-1} were single experiments.

Cultivations were operated with a ~300ml working volume in 1.25 L Biobundle bioreactors (Applikon Biotechnology B.V., Schiedam, the Netherlands) controlled either by an ADI 1030 or ez-Control biocontroller (Applikon Biotechnology B.V.) and either a cultivation control program "BioXpert NT" or "BioXpert XP" (Applikon Biotechnology B.V.). The system was equipped with OD, pH, pO_2, off-gas CO_2 and O_2 and temperature sensors.

Growth characteristics were calculated on the basis of total volume of medium pumped out from bioreactor (L), biomass (gram dry cellular weight (gDCW)/L), organic acid and glucose concentrations in culture broth and medium (mM), and CO_2 and O_2 concentrations in the outflow gas (mM). Detailed description of A-stat and D-stat cultivation systems together with control algorithms and formulas used to calculate growth characteristics can be seen in [137].

D-stat experiments in this thesis were carried out slightly differently from the classical D-stat [137] by using two growth media (the main and the second medium; see above) instead of smoothly changing an environmental parameter. Acetate consumption capability was studied in these D-stats by increasing the feed rate of the glucose and acetate-supplemented medium (second medium) while decreasing that of the glucose-supplemented medium (main medium) after steady state had been achieved in chemostat on the latter medium. This scheme maintained a constant D and glucose concentration in the total feed.

3.4. Analytical methods

3.4.1. Biomass concentration and cell count determination

Biomass concentration in culture broth (X) expressed as gDCW/L was determined gravimetrically either with the filtration or centrifugation method described in detail in Publication III and [6], respectively. Coefficient of variation (CV) for the determination of X with these methods between technical replicates was <2%. Simultaneous measurement of X and optical density (at 600 nm) of culture broth (OD) was done with high resolution of specific growth rate (μ) to determine a reliable correlation factor (K; K = X / OD) dependence on OD and μ, so that X could be accurately calculated also for μ values where X was not directly measured.

Cell counts were measured by incubating five replicate LB-agar plates at 37 °C for ~11–12 h and expressed as CFU ml^{-1}, equivalent to cell ml^{-1}.

3.4.2. Exo-metabolome analysis

Samples of culture broth were centrifuged (14,000 × g for 5 min), and the supernatant was collected and analyzed for glucose and organic acids by HPLC (Alliance 2795 system, Waters Corporation, Milford, MA, USA) using either a Bio-Rad HPX-87H Aminex ion-exclusion (Bio-Rad Laboratories, Inc., Hercules, CA, USA) or Agilent Hi-Plex H (Agilent Technologies, Santa Clara, CA, USA) column connected to RI and UV detectors (35 °C, flow rate 0.6 mL/min). The column was eluted with 4.1 mM sulphuric acid for glucose, carbamoyl-aspartate (CBASP), lactate and orotate and with 26.5 mM formic acid for acetate, dihydroorotate (DHO) and acetyl-aspartate (NAA) analysis.

3.4.3. Transcriptomics

Genome-wide transcriptome analysis for 4,321 transcripts was conducted in one A-stat experiment with six technical replicates for reference sample at $\mu = 0.11$ h^{-1} using Agilent DNA microarrays producing gene expression ratios between $\mu = 0.21$; 0.26; 0.31; 0.36; 0.40; 0.48 h^{-1} and 0.11 h^{-1} (chemostat point prior to the A-stat phase) (GEO reference series: GSE23920). In short: after sampling, RNA degradation was halted, total RNA extracted, cDNA synthesized and labeled and hybridized; slides were scanned and data was processed in the R environment using global lowess normalization.

Absolute quantification of transcriptomes (molecules per cell or per fL of biomass) at $\mu = 0.11$; 0.21; 0.31; 0.40; 0.48 h^{-1} was performed on the basis that oligo spot intensities of the Agilent platform can be used as a proxy for mRNA abundances, since spot intensities and mRNA abundances correlate perfectly ($R^2 = 1.00$; see Figure 6 in Agilent Application Note 5989-9159EN). In short: average spot intensities were summed corresponding to the total amount of mRNA which was converted to unit g total mRNA per gDCW by determining μ-dependent total RNA% in DCW (Table S1 in Publication IV) using the RNeasy Mini Kit (QIAGEN, Valencia, CA, USA) and assuming mRNA content in total RNA to be 5% [138,139]; mRNA molecule numbers in gDCW were calculated by the fraction of each mRNA's spot intensity from the summed spot intensities taking into account the molecular weight of each mRNA. Finally, mRNA abundances (molecules per cell) and concentrations (molecules per fL of biomass) were calculated from the previous values based on determined μ-dependent biomass concentration and cell counts in the culture broth, 30% dry fraction of wet biomass [140–142] and cell buoyant density of 1 g ml^{-1} to estimate cell volume.

3.4.4. Proteomics

Proteome analysis in Publication I based on ^{15}N-labelling and LC-MS/MS was conducted in two independent A-stat experiments producing protein expression ratios for around 1,600 proteins between $\mu = 0.20 \pm 0.01$; 0.26; 0.30 ± 0.01; 0.40 ± 0.00; 0.49 ± 0.01 h^{-1} and 0.10 ± 0.01 h^{-1} (chemostat point prior to the A-stat phase) (PRIDE

accession numbers 12189–12199). In short: samples were flash frozen in liquid nitrogen, total protein extracted, sample at each μ was pooled together in equal amounts with ^{15}N-labelled *E. coli* biomass ('internal standard') and separated on gel followed by protein digestion with trypsin, peptide extraction and LC-MS/MS analysis. Protein expression ratios between A-stat samples and chemostat reference point at $\mu = 0.10 \pm 0.01$ h^{-1} were calculated through ratios between each cultivation sample and the 'internal standard'. Proteins were identified and quantified for each μ at a >95% confidence interval in average from 89,303 distinct two or more high-confidence peptides with a Mascot score >20.

Absolute quantification of proteomes (molecules per cell or per fL of biomass) was performed in two independent A-stat experiments for 1,185 proteins in chemostat reference points at $\mu = 0.11$ h^{-1} using the label-free iBAQ quantification approach [99] and published in [143]. In short: a cultivation sample was pooled together with the Universal Proteomics Standard (UPS2; Sigma-Aldrich), which is a mixture of proteins with known amounts, differing in size and concentrations, followed by protein digestion, extraction and LC-MS/MS analysis. Protein concentrations in chemostat samples were determined using precursor ion current areas and UPS2 standard curve. Finally, intracellular protein abundances (molecules per cell) and concentrations (molecules per fL of biomass) were calculated from the previous values based on determined μ-dependent biomass concentration and cell counts in the culture broth, 30% dry fraction of wet biomass [140–142] and cell buoyant density of 1 g ml^{-1} to estimate cell volume.

Absolute quantification of proteomes (molecules per cell or per fL of biomass) in Publication IV for $\mu = 0.20$; 0.30; 0.40; 0.49 h^{-1} was carried out based on absolute quantification at $\mu = 0.11$ h^{-1} and relative protein expression data described above based on determined μ-dependent biomass concentration, cell counts and total protein % in DCW (Table S1 in Publication IV) measured by the Lowry method [144], 30% dry fraction of wet biomass [140–142] and cell buoyant density of 1 g ml^{-1} to estimate cell volume.

Details of proteome analysis conducted in Publication III can be found in the latter publication.

3.5. Metabolic flux analysis

Metabolic flux analysis (MFA) was conducted using a simplified metabolic network taking into account the central carbon metabolism pathways—glycolysis, pentose phosphate pathway (PPP), tricarboxylic acid (TCA) cycle—, a part of the pyrimidine pathway (to include CBASP, DHO and orotate) and the NAA synthesis reaction (Figure S1 in Publication II). The reconstructed network converted into a fully determined and calculable stoichiometric matrix consisting of 24 metabolites and 50 fluxes (24 unknown, 1 measured inflow, 7 outflow and 18 calculated fluxes based on biomass composition and stoichiometries of anabolic pathways) taking into account ATP, NADH and NADPH stoichiometry. A constant P/O = 2 not dependent on μ was assumed in the calculations. We also determined μ-dependent biomass composition:

RNA as described above; amino acids with UPLC [145]; lipids with an UPLC-MS method [51] and polysaccharides based on lipids data (Table S1 in Publication II). DNA content was taken from [55], ash and residual water in dry biomass were calculated as residual of the latter. The only difference between MFA conducted in Publications I and III was that oxaloacetate or pyruvate, respectively, was set as the missing carbon outflow flux not identified experimentally. Refer to Additional file 2 in Publication II for MFA details.

3.6. Covariance analysis

Apparent *in vivo* catalytic rates of enzymes (k_{app}, s^{-1}) were calculated for 191 metabolic enzymes associated with catalyzing the fluxes calculated by MFA and 52 ribosomal proteins assuming each protein chain being catalytically active. Proteins were assigned to fluxes according to the EcoCyc database [64]. The sum of all amino acid synthesis fluxes was used as the flux catalyzed by ribosomal proteins. k_{app} values were calculated as follows:

$$k_{app_i} = \frac{specific\ flux\ rate_i}{protein_i} \times \frac{N_A \times 0.3}{10^{12}}$$

where specific flux $rate_i$ (mol/gDCW/s) is the flux catalyzed by $protein_i$ (molecules/fL of biomass), N_A is the Avogadro number (~6.02×10^{23}), 0.3 is dry fraction of one gram of wet biomass [140–142] and 10^{12} is the conversion factor from fL to g assuming a buoyant density of 1 g ml^{-1}.

We used covariance analysis to determine the relevant most and least costly proteins using protein synthesis cost (nATP; see below for calculation) over the whole range of μ and for statistically determining the regulation levels (*e.g.* transcriptional (TR), translational (TL)) as it describes both the direction and magnitude of mRNA, protein and flux changes with increasing μ making it a suitable statistical method for analysis of absolute quantitative data.

Covariance (COV) was calculated according to the following formulas:

for nATP as:

$$COV_{nATP} = \frac{1}{n-1}\sum_{i=1}^{n}(nATP_i - \overline{nATP})(\mu_i - \bar{\mu})$$

for gene expression regulation analysis at protein and mRNA level (pm) as:

$$COV_{pm} = \frac{1}{n-1}\sum_{i=1}^{n}\left(prot_i/mRNA_i - \overline{prot/mRNA}\right)(\mu_i - \bar{\mu})$$

for gene expression regulation analysis at specific flux rate and protein level (k_{app}) as:

$$COV_{k_{app}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(spec\ flux_i/prot_i - \overline{spec\ flux/prot}\right)(\mu_i - \bar{\mu})$$

where mRNA, prot and spec flux represent mRNA and protein concentrations, and specific flux rates measured at their respective μ.

nATP was calculated according to formula:

$$x_{nATP} = x_{prot} \times (x_{nAA} - 1) \times 4.306$$

where x_{prot} is the protein concentration measured at the respective μ, x_{nAA} is the number of amino acids in the protein and 4.306 represents the cost of polymerization of one amino acid into the growing peptide chain by the ribosome in ATP [55].

Refer to Publication IV for details of covariance analysis. In short: in addition to determination of COV values, uncertainty values of COV were calculated for the statistical hypothesis testing using one-sided Z-test of covariance values being statistically different or not different from zero. The significance threshold was set to 0.05 and tests were also subjected to false discovery rate (FDR) filtering at level $\alpha = 0.05$ according to the Benjamini–Hochberg method [146].

According to the results of Z-test, genes were divided into three groups. One group corresponded to genes with a covariance value statistically higher than zero. In gene expression regulation analysis, these genes' expression regulation level is referred to as post-TR or TL for control of protein levels, and post-TL for control of flux levels. The second group corresponded to genes with a covariance value equal to zero at the statistically significant level. In gene expression regulation analysis, these genes' expression regulation level is referred to as TR for control of protein levels, and TL for control of flux levels. The rest of the genes (not included in the first two groups) were described by such a high uncertainty level of covariance that it was impossible to determine their nature towards zero.

4. RESULTS AND DISCUSSION

The results of this systems biology study of specific growth rate (μ)-dependent metabolism of *Escherichia coli* are presented and discussed in the following six sections organized by topic.

4.1. Macroscopic growth characteristics (Publications I and II)

4.1.1. Characterization of metabolic switch-points

We chose to analyze μ-dependent metabolism of *E. coli* by cultivating the cells in accelerostat (A-stat) cultures [3] which allows to study the cells in strictly defined physiological states determined by controlled growth conditions and describes steady state physiology equally to chemostat [20–26]. Moreover, A-stat enables to collect vast amount of data in short time to monitor the dynamics of metabolism with very high resolution and precisely detect metabolic switch-points (*e.g.* start of acetate overflow).

E. coli K-12 MG1655 was grown in three independent A-stat experiments on defined minimal medium supplemented with 4.5 g/L glucose under aerobic conditions at 37 ^{0}C and pH 7. Glucose-limited cultures were first stabilized in chemostat at dilution rate (D) = 0.1 h^{-1} to achieve steady state conditions after which the A-stat phase was started with a continuous increase of D using acceleration 0.01 h^{-2}. Continuous increase of μ (μ = D in steady state) enabled to detect several important changes in *E. coli* metabolism until cells could not keep up with the rising D after achieving maximum specific growth rate (μ_{max}) at 0.54 ± 0.03 h^{-1} (average ± standard deviation; n = 3), resulting in culture washout (Figure 1).

Cells reached maximum specific CO_2 production (q_{CO2}) and O_2 consumption rates (q_{O2}) at μ = 0.46 ± 0.02 h^{-1} after which metabolic fluctuations were observed until μ_{max}. The nature of these fluctuations was not studied in much detail but simultaneous and inverse fluctuations to CO_2 in residual glucose levels were detected as well. We were successful in precisely detecting the metabolic switch-point of metabolism shifting from fermentative to respire-fermentative growth—start of acetate overflow metabolism—at μ = 0.27 ± 0.02 h^{-1} (see arrow in Figure 1). Interestingly, the high-resolution A-stat data also detected a two-phase acetate accumulation profile: slow accumulation started at μ = 0.27 ± 0.02 h^{-1} with concomitant change in q_{CO2} while faster accumulation of acetate commenced after cells had reached maximum q_{CO2} at μ = 0.46 ± 0.02 h^{-1}. The latter correlation between acetate and CO_2 fluxes can be explained by the fact that less acetyl-CoA is available to enter the tricarboxylic acid (TCA) cycle if carbon is lost into acetate. Notably, the faster accumulation phase of acetate was preceded by a sharp decline of the important carbon catabolite repression (CCR) signaling molecule cyclic AMP (cAMP) (Figure 1). It was interesting to detect the dynamics of cAMP with rising μ together with the acetate profile as cAMP could play a role in the regulation of acetate overflow (see *Acetate overflow metabolism* for details).

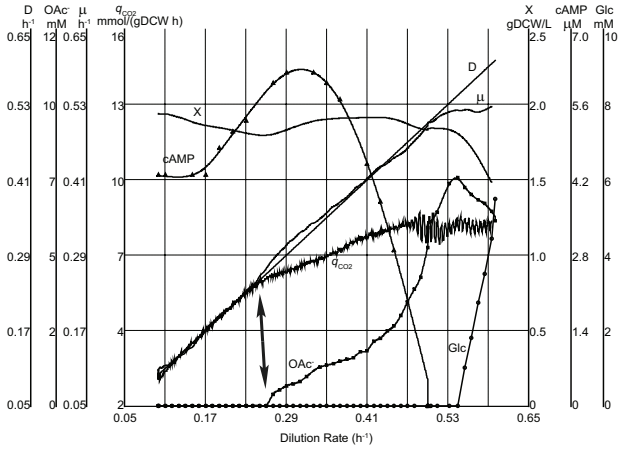

Figure 1. Specific growth rate-dependent E. coli K-12 MG1655 metabolism characterization in one A-stat cultivation. *D, dilution rate; X, biomass concentration (gram dry cellular weight (gDCW)/L); μ, specific growth rate; q_{CO2}, specific CO_2 production rate; OAc^-, acetate concentration; Glc, glucose concentration; cAMP, cyclic AMP concentration. Arrow indicates the start of acetate overflow. Start of vertical axes was chosen for better visualization.*

As expected, the specific glucose consumption rate (q_{glc}) increased proportionally with rising μ (R^2 = 0.99) from 1.47 ± 0.07 to 6.35 ± 0.13 mmol/(gram dry cellular weight (gDCW) h) from μ = 0.11 to 0.47 h^{-1}, respectively. This also means that the biomass yield on glucose (Y_{XS}) was constant at 0.41 ± 0.01 gDCW/g glucose throughout the studied range of μ.

4.1.2. Detailed analysis of exo-metabolome and carbon balance

We noted that although the total carbon flow from glucose to the 'well-known' products of biomass, CO_2, acetate and lactate almost closed the carbon balance at slow growth (~97%), their fraction from the carbon balance decreased with rising μ and covered only ~82% near μ_{max}. As there were no indications that the measurement accuracy of these compounds could worsen with increasing μ, the latter phenomenon hinted that cells might increasingly divert carbon to other by-products with faster growth. Indeed, evidence for this came from carefully studying the HPLC chromatograms which revealed several unknown peaks increasing at faster growth. We were able to identify these peaks using HPLC-MS as orotate, dihydroorotate (DHO), carbamoyl-aspartate (CBASP) and acetyl-aspartate (NAA) based on the m/z values observed for each peak and literature survey [36,37].

Accumulation of the pyrimidine pathway compounds—orotate, DHO and CBASP—can be explained by the *E. coli* K-12 MG1655 genotype. This specific strain is prone to pyrimidine starvation due to an *rph* frameshift mutation leading to low *pyrE* (encodes PyrE protein which catalyzes orotate conversion into orotidine-5-phosphate) expression [147], which could possibly lead to accumulation of pyrimidine pathway precursor molecules which all the latterly mentioned compounds are (Figure 2). This

was proven by chemostat experiments using medium supplemented with uracil as it almost completely abolished accumulation of pyrimidine pathway intermediates (data not shown).

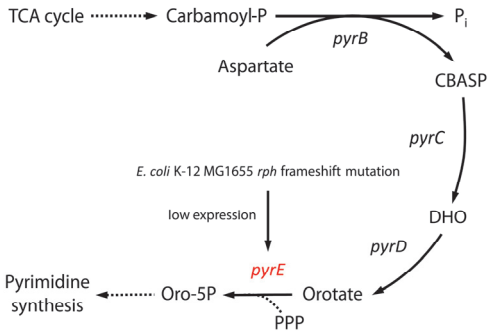

Figure 2. rph frameshift mutation triggered accumulation of pyrimidine pathway precursor compounds in E. coli K-12 MG1655. *Carbamoyl-P, carbamoyl-phosphate; CBASP, carbamoyl-aspartate; DHO, dihydroorotate; Oro-5P, orotidine-5-phosphate; TCA cycle, tricarboxylic acid cycle; PPP, pentose phosphate pathway; pyrB, aspartate carbamoyltransferase; pyrC, dihydro-orotase; pyrD, dihydro-orotate oxidase; pyrE, orotate phosphoribosyltransferase. Gene names are in italic.*

Although excretion of these pyrimidine pathway intermediates by *E. coli* has been observed before [4,36,37], these carbon wasting outflows are generally not taken into account in metabolic flux analysis (MFA) studies. Thus the metabolic network with the missing outflows is not completely accurate possibly leading to questionable conclusions. Therefore, determination of μ-dependent carbon wasting profiles of these pyrimidine pathway intermediates is important for more accurate description of metabolic regulation.

Excretion of pyrimidine pathway intermediates during increase of μ was detected in three phases (Figure 3). DHO and CBASP accumulated increasingly up to the start of acetate overflow. After this, DHO started to decline whereas orotate and CBASP leveled off until their levels started to rise again simultaneously (Figure 3) with the sharp decrease of cAMP and faster accumulation of acetate (Figure 1). This rise could be explained by the high demand for RNA synthesis at higher μ which leads to precursor molecule accumulation because of the low *pyrE* expression. These observations demonstrate a strong link between acetate overflow and carbon wasting into other products. It seems probable that this CBASP-DHO-orotate bottleneck can lead to the excretion of NAA since oxaloacetate is over-produced in the TCA cycle and this excess carbon cannot be shunted towards the pentose phosphate pathway (PPP).

Quantification of these additional carbon wasting outflows increased the accuracy of MFA (see *Flux dynamics*) and a possible limitation of μ_{max} by RNA synthesis through precursor accumulation and carbon wasting due to the *E. coli* K-12 MG1655 genotype proposes a way how to increase μ_{max} and Y_{XS} which is relevant for the biotechnology industry.

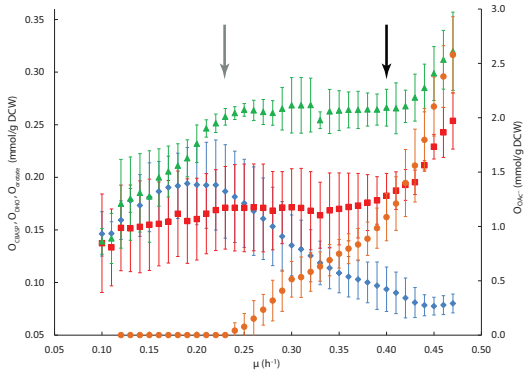

Figure 3. Specific growth rate-dependent dynamic carbon wasting profiles in three E. coli K-12 MG1655 A-stat cultivations. *μ, specific growth rate. O, production of compound per biomass (mmol/gram dry cellular weight (gDCW)): CBASP, carbamoyl-aspartate (green triangle); DHO, dihydroorotate (blue diamond); orotate (red square); OAc^-, acetate (orange circle). Grey arrow denotes acetate overflow switch with concomitant stop of DHO and CBASP increase whereas black arrow depicts faster acetate accumulation coupled induction of orotate and CBASP excretion. Error bars represent standard deviation of triplicate A-stat experiments.*

Detailed carbon balance analysis in A-stat showed that 47.2 ± 0.8% of carbon was used for biomass synthesis throughout the studied range of μ while the fraction of carbon used for CO_2 production decreased with rising μ from ~49% to ~29% (Figure 4). Importantly, inclusion of the previously non-accounted by-products orotate, DHO, CBASP and NAA clearly improved the carbon balance throughout the studied range of μ as the carbon balance became fully closed at slowest and reached ~87% at faster growth. The latter still points to loss of carbon into some other not detected compounds. In detail, total carbon wasting into by-products (excluding CO_2) increased from 3 to 11% from the carbon balance comparing μ = 0.11 to 0.47 h^{-1} (Figure 5). Acetate quickly became the main excreted compound after the start of acetate overflow.

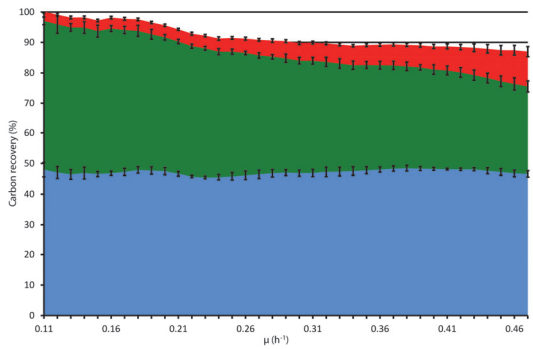

Figure 4. Specific growth rate-dependent carbon balance in three E. coli K-12 MG1655 A-stat cultivations. *μ, specific growth rate. Carbon recovery % in carbon balance into: biomass (blue); CO_2 (green); sum of carbon wasting into acetate, lactate, orotate, carbamoyl-aspartate, dihydroorotate and acetyl-aspartate (red). Error bars represent standard deviation of triplicate A-stat experiments.*

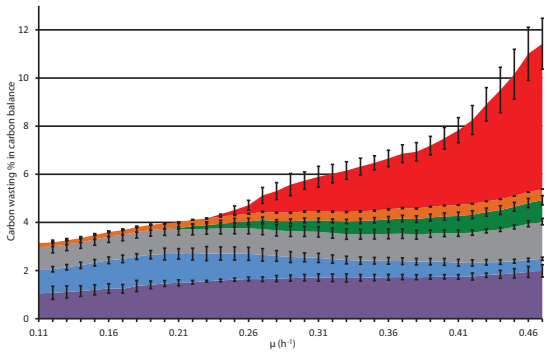

Figure 5. Specific growth rate-dependent dynamic carbon wasting profiles in three E. coli K-12 MG1655 A-stat cultivations. *μ, specific growth rate. Carbon wasting % in carbon balance into: acetate (red); lactate (orange); acetyl-aspartate (green); orotate (grey); dihydroorotate (blue); carbamoyl-aspartate (violet). Error bars represent standard deviation of triplicate A-stat experiments.*

4.1.3. A-stat reproducibility and comparison with chemostat

Our three A-stat cultures were highly reproducible, evidenced by the less than 10% variation among the detected switch-points and other macroscopic growth characteristics highlighted above (Table 1). Also the protein concentrations and their μ-dependent dynamics in two independent A-stat cultures showed high reproducibility (see below).

Table 1. A-stat reproducibility and comparison of A-stat and chemostat growth characteristics

	μ = 0.24 h^{-1}		μ = 0.30 h^{-1}		μ = 0.40 h^{-1}		μ = 0.51 h^{-1}		μ = 0.10-0.47 h^{-1}
	Chemostat	A-stat	Chemostat	A-stat	Chemostat	A-stat	Chemostat	A-stat	A-stat RSD, %
Y_{XS} [a]	0.44	0.40 ± 0.01	0.46	0.41 ± 0.01	0.44	0.42 ± 0.00	0.43	0.41 ± 0.01	2.0
Y_{OAc-} [b]	NDE	NDE	0.53	0.90 ± 0.32	1.70	1.56 ± 0.23	3.25	3.35 ± 0.82	ND
Y_{cAMP} [c]	3.47	3.59 ± 0.39	3.25	3.55 ± 0.32	2.70	2.17 ± 0.07	0.86	0.71 [e]	9.1
Y_{CO2} [d]	27.56	30.12 ± 2.04	27.55	27.19 ± 1.22	26.24	23.86 ± 1.41	ND	21.19 ± 0.19	5.6

A-stat values represent the average of three independent experiments and standard deviation follows the ± sign. Chemostat values from one experiment. μ, specific growth rate. NDE, not detected. ND, not determined. RSD, relative standard deviation. [a], biomass yield (gram dry cellular weight (gDCW)/g glucose). [b], acetate production per biomass (mmol/gDCW). [c], cAMP production per biomass (mmol/gDCW). [d], CO_2 production per biomass (mmol/gDCW).[e], data from one A-stat experiment.

An important parameter in A-stat cultivations is the selected acceleration of D. If the appropriate value is chosen, A-stat data could be considered to describe steady state physiology equally to chemostat. The acceleration 0.01 h^{-2} used in all A-stats reported in this thesis yielded quantitatively comparable results with chemostat cultures at the level of both macroscopic growth characteristics (Table 1 & Table 2), and transcriptome and proteome levels (Figure S3 in Publication I and [25]). Thus all the quasi steady state data of A-stat covered in this thesis can be considered to represent

steady state physiology similar to chemostats, as also concluded before for different organisms if the appropriate acceleration is chosen [20–26].

Table 2. Comparison of carbon wasting in A-stat and chemostat

	μ = 0.10 h^{-1}		μ = 0.24 h^{-1}		μ = 0.30 h^{-1}		μ = 0.45 h^{-1}	
	Chemostat	A-stat	Chemostat	A-stat	Chemostat	A-stat	Chemostat	A-stat
O_{CBASP}	0.188	0.139 ± 0.047	0.089	0.218 ± 0.018	0.195	0.268 ± 0.026	0.188	0.299 ± 0.026
O_{DHO}	0.131	0.146 ± 0.021	0.117	0.181 ± 0.043	0.091	0.135 ± 0.029	0.065	0.078 ± 0.008
$O_{lactate}$	ND	0.021 ± 0.005	0.034	0.041 ± 0.003	0.034	0.049 ± 0.006	0.046	0.059 ± 0.002
O_{NAA}	ND	ND	0.006	ND	0.034	0.057 ± 0.013	0.069	0.111 ± 0.021
$O_{orotate}$	0.091	0.137 ± 0.047	0.193	0.171 ± 0.041	0.115	0.171 ± 0.037	0.344	0.229 ± 0.012

μ, specific growth rate. A-stat values represent the average of three independent experiments and standard deviation follows the ± sign. Chemostat values from one experiment. ND, not detected. O, production of compound per biomass (mmol/gram dry cellular weight (gDCW)): CBASP, carbamoyl-aspartate; DHO, dihydroorotate; NAA, acetyl-aspartate.

4.2. Metabolic flux dynamics and energy metabolism (Publication II)

To gain insights into μ-dependent regulation of metabolic flux dynamics and energy metabolism in *E. coli*, we conducted a simplified MFA to map carbon flows through central carbon metabolism (CCM) with two advantages over classical MFA studies. Firstly, we included the novel carbon wasting routes in our MFA obtained by detailed carbon balance analysis which makes our MFA more accurate as the missing carbon is usually accounted as a flux to a single excreted compound. Although the absolute amount of these excreted substances in the carbon balance is not substantial (less than 5%), linking their accumulation dynamics to μ and metabolic routes is relevant for acknowledging the potential imbalance between pyrimidine metabolism, acetate recycling and non-growth associated ATP production (see below). If orotate, DHO, CBASP and NAA by-product outflows would be excluded from MFA and missing carbon accounted solely as pyruvate outflow, pyruvate dehydrogenase (PDH), TCA cycle and PEP carboxylase fluxes would deviate by 11, 24 and 60%, respectively, at μ = 0.47 h^{-1} from the values calculated by our model (Table S6 in Publication II). Secondly, rarely is biomass composition experimentally determined in MFA studies but taken from the literature [4,9,27,39]. Furthermore, dependence of biomass composition on growth conditions [27,39,51–53] is also neglected, and both of the latter could possibly lead to distorted MFA results and equivocal conclusions. Therefore, we experimentally determined μ-dependent biomass composition of amino acids, RNA, lipids and polysaccharides (Table S1 in Publication II). Inclusion of μ-dependent biomass composition in MFA was important as highlighted by up to a 15% difference in flux values calculated using either μ-dependent or constant biomass composition (Table S2 in Publication II), also shown before by sensitivity analysis [52].

Our simplified fully determined metabolic network (Figure S1 in Publication II) consisted of three main CCM pathways—glycolysis, PPP, TCA cycle—, a part of the pyrimidine pathway (to include CBASP, DHO, orotate) and the NAA synthesis reaction with 50 fluxes and 24 metabolites taking into account ATP, NADH and NADPH stoichiometry.

4.2.1. Metabolic flux dynamics

All specific flux rates increased with faster growth. However, to have a better description of changes in carbon flow distribution within the metabolic network, we compared fluxes in units of flux per biomass (mmol/gDCW). In addition to quantitatively describing the flux magnitude per cell biomass, it also characterizes the carbon flow through the flux relative to total consumed carbon in these experiments since Y_{XS} (gDCW/g glucose) varied only 2% over the range of μ *i.e.* flux per biomass is equivalent to flux per consumed carbon.

It is clearly seen from MFA that acetate overflow plays an important role in overall flux patterns since major carbon flux re-distribution was initiated with the start of acetate overflow at $\mu = 0.27 \pm 0.02$ h^{-1} with simultaneous change in q_{CO2} (Figure 1). Acetate overflow reduced the carbon flow from glycolysis and acetate recycling in the phosphotransacetylase-acetyl-CoA synthetase (PTA-ACS) node (see *Importance of PTA-ACS node*) to acetyl-CoA and CCM, triggering the reduction of TCA cycle fluxes (Figure 6). Decrease of TCA cycle throughput was also seen by the decline of the proportion of CO_2 and NADH produced by the TCA cycle (Additional file 3 in Publication II). Start of acetate overflow and reduction of TCA cycle throughput subsequently led to the induction of PPP fluxes, reduction of glycolysis (Figure 6) and ATP produced from it (Figure 8), most probably to meet the increasing NADPH demands for higher RNA synthesis with faster growth [55–58]. Although the above-mentioned *E. coli* μ-dependent patterns of TCA cycle, glycolysis and PPP have also been observed in general using glucose-limited chemostat cultures and ^{13}C-MFA analysis [4,7–9], the switch-point in flux dynamics can exactly be captured only with A-stat cultures.

An important branch point flux in CCM—PDH—reached its maximum throughput at $\mu = 0.42$ h^{-1} with concomitant slight increase in glycolysis fluxes also resulting in accelerated carbon wasting into by-products (Figure 3 & Figure 6). In addition, decrease of flux through pyruvate kinase (Pyk) with increasing PEP carboxylase (Ppc) and Vprod (missing carbon outflow flux from oxaloacetate in our model) fluxes during slow growth until the start of acetate overflow shows that some of the consumed carbon was in excess and excreted as oxaloacetate through the Ppc flux (Figure 6). These observations demonstrate a strong link between acetate overflow and carbon wasting into other products.

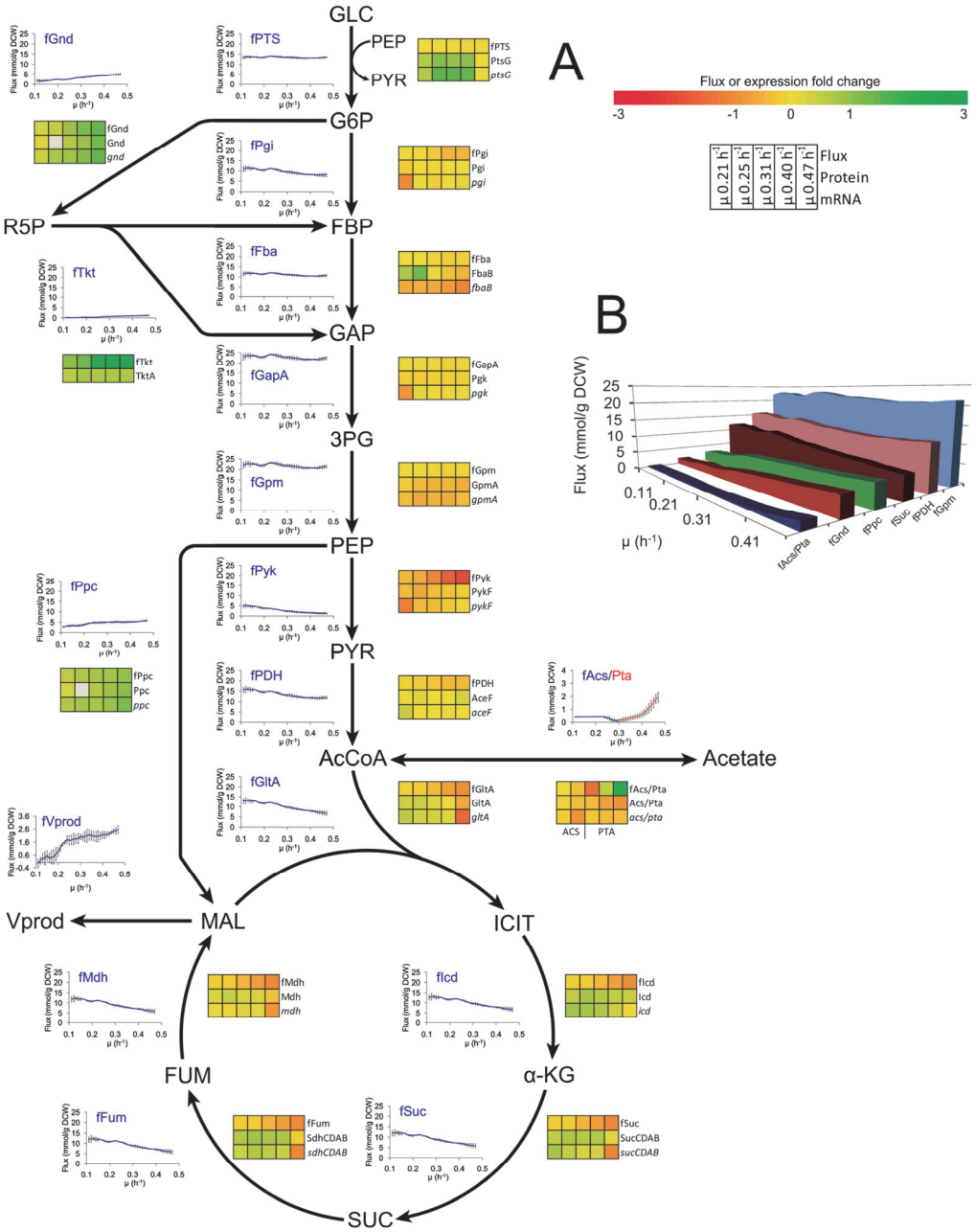

Figure 6. Metabolic flux, protein and mRNA dynamics in E. coli K-12 MG1655 central carbon metabolism with rising specific growth rate in three A-stat cultivations. *μ, specific growth rate.* ***A.*** *Flux or expression fold change (log2) is calculated for each μ compared to μ = 0.10 h^{-1}; grey box depicts missing value; Acs flux was switched to Pta after μ = 0.31 h^{-1} in MFA since acetate excretion exceeds its production accompanying biosynthesis. Error bars represent standard deviation of triplicate A-stat experiments.* ***B.*** *Selected glycolysis, TCA cycle, PPP, anaplerotic and acetate related fluxes.*

4.2.2. Energy metabolism

The part of unaccounted ATP in MFA calculations for the analysis of energy metabolism is always an intriguing aspect. It is important to point out that we refer to this imbalance between ATP production in catabolism and consumption for anabolism as non-growth associated ATP production (ATP spilling; mmol of ATP/gDCW). We use ATP spilling and avoid using terms maintenance or non-growth associated and growth-associated maintenance due to the confusion regarding their exact biological nature [44,45,60].

It became clear from MFA calculations that acetate overflow also plays an important role in energy metabolism as disruption of the ATP-wasting PTA-ACS node resulting in acetate overflow at $\mu = 0.27 \pm 0.02$ h^{-1} reduced ATP spilling by 36% with rising μ (Figure 7). This response in energy metabolism was detected in this study since the futile PTA-ACS cycle (an equivalent amount of ATP to acetate is wasted with acetate recycling) shown to operate *in vivo* in Publications I and III (see below for details) and in [33,34] was included into the model network.

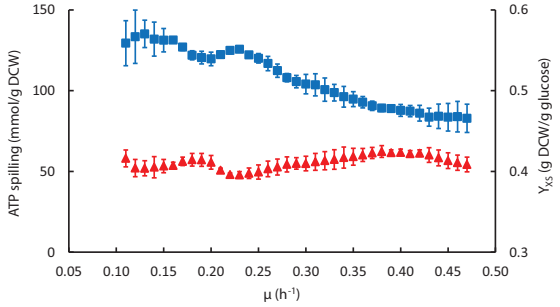

Figure 7. Specific growth rate-dependent ATP spilling and biomass yield in three E. coli K-12 MG1655 A-stat cultivations. *μ, specific growth rate; gDCW, gram dry cellular weight; ATP spilling, non-growth associated ATP production (blue square); Y_{XS}, biomass yield (red triangle). Error bars represent standard deviation of triplicate A-stat experiments.*

Decrease of ATP spilling (40 mmol/gDCW in total) might indicate increase of Y_{XS}, however, it remained constant at 0.41 ± 0.01 gDCW/g glucose throughout the studied range of μ (Figure 7). This apparent discrepancy between the decrease in ATP spilling and constant Y_{XS} (Figure 7) could be explained by the fact that carbon wasting increased from 3 to 11% with rising μ (Figure 5) as follows. As acetate recycling by synthesis and assimilation in the PTA-ACS node (Publication I and III) is a futile cycle, an equivalent amount of ATP to acetate is simultaneously wasted with synthesis and assimilation of acetate (see *Importance of PTA-ACS node* for the explanation of its biological relevance). Hence, accumulation of acetate likely triggers a 36% decline of ATP spilling (Figure 7) since re-assimilation of acetate (wasting 1 molecule of ATP) decreases with rising μ after the start of acetate overflow. This energy save is, however, counteracted by the increase of carbon wasting in the carbon balance from 3 to 11%, which results in constant Y_{XS}. However, *E. coli* might also possess additional mechanisms to maintain constant Y_{XS} under increasing carbon wasting conditions.

Expectedly, the total specific ATP production rate (q_{ATP}) increased with rising μ but changed its slope after the start of acetate overflow (Figure 8). The non-growth associated maintenance cost is assumed to be constant and generally estimated by extrapolating q_{glc} measured in chemostat cultures to $\mu = 0\ h^{-1}$ [9,40,44,45,60], based on the belief that energy production or substrate consumption is linear with μ. However, our A-stat data (Figure 8) clearly demonstrates that this 'extrapolation' method is not accurate. This is because the non-growth associated maintenance cost estimated as explained above using data points from before or after the start of acetate overflow $\mu = 0.27 \pm 0.02\ h^{-1}$ yield very different values of ~3 and ~19 mmol/(gDCW h), respectively. Although these values are within the estimates found in the literature [10,39,44,45,60,148], the 6-fold difference between the estimated values demonstrates that even high-resolution data of A-stat should not be used for estimating non-growth associated maintenance costs using the 'extrapolation' method.

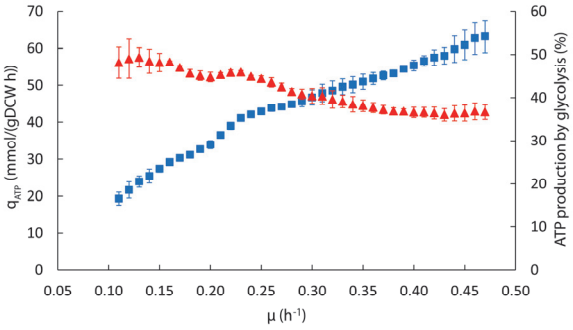

Figure 8. Specific growth rate-dependent total specific ATP production rate and proportion of ATP production by glycolysis in three E. coli K-12 MG1655 A-stat cultivations. *μ, specific growth rate; gDCW, gram dry cellular weight; q_{ATP}, total specific ATP production rate (blue squares); ATP production by glycolysis (red triangle). Error bars represent standard deviation of triplicate A-stat experiments.*

4.3. Transcriptome and proteome dynamics (Publication I)

Adaptation of microorganisms to different environmental conditions is usually accompanied by a change in μ and of many properties of cell physiology *e.g.* cell morphology [1–3], gene expression [4–8], metabolic flux patterns [4,7–10]. The majority of these mechanisms are regulated through changes in gene expression. Hence, to gain detailed insights into μ-dependent regulation of gene expression, we carried out simultaneous functional-genomics analyses of transcriptome and proteome in the same A-stat cultures described above.

DNA microarray analysis of 4,321 transcripts was conducted with the Agilent platform using samples from one A-stat cultivation. Gene expression ratios between $\mu = 0.21$; 0.26; 0.31; 0.36; 0.40; 0.48 h^{-1} and 0.11 h^{-1} (chemostat point prior to the A-stat phase) were calculated. Using a ^{15}N-labelling and LC-MS proteome approach, *E. coli* protein expression ratios for around 1,600 proteins were generated by comparing two independent A-stat cultures at $\mu = 0.20 \pm 0.01$; 0.26; 0.30 ± 0.01; 0.40 ± 0.00; $0.49 \pm$

0.01 h^{-1} with 0.10 ± 0.01 h^{-1} (chemostat point). High reproducibility of both A-stat cultures and our proteome approach is demonstrated by high correlations of R = 0.8–0.9 (Pearson correlation coefficient) between protein expression changes in two independent cultures (Figure 9A).

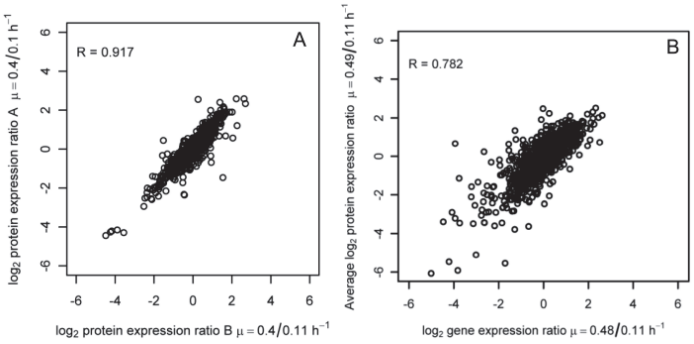

Figure 9. Reproducibility of proteome dynamics, and correlation of transcriptome and proteome changes in E. coli K-12 MG1655 A-stat experiments. *μ, specific growth rate. R, Pearson correlation coefficient. Data for all μ is in Additional file 1 in Publication I.* ***A.*** *Correlation between log2 protein expression changes between μ = 0.4 and 0.1 h^{-1} in two independent A-stats. A and B represent duplicate experiments.* ***B.*** *Correlation between log2 mRNA and protein expression changes between μ = 0.4 and 0.1 h^{-1}. Average protein expression ratio is from two independent A-stats. Genes with coefficient of variation smaller than 20% among its multiple spots on DNA microarrays were included for analysis.*

4.3.1. Correlation of transcriptome and proteome dynamics

Analyzing the correlation between changes in mRNA (transcriptome) and protein (proteome) levels can point to levels of gene expression regulation (*e.g.* transcriptional, post-transcriptional). Several recent reviews [67–70] have highlighted that time-course analyses of bacteria and yeast reveal large differences between mRNA and protein abundance changes in perturbed systems indicating considerable post-transcriptional regulation. Hence, they conclude that understanding of perturbed systems is still incomplete and requires further analysis. Interestingly, our perturbed *E. coli* cultures show high correlations up to R = 0.8 between mRNA and protein changes with rising μ (Figure 9B).

Investigating the effects of –omics data processing to correlations revealed that the initially observed correlations (R up to 0.8) are maximum for our data since neither increasing threshold limits in mRNA and protein measurements, nonlinear transformation nor removing outliers improved correlations (data not shown). As cells in the studies analyzed by the reviews [67–70] were mostly cultured in non-steady state conditions, the high mRNA-protein correlations observed in our dynamic experiments conducted under strictly defined and controlled growth conditions might imply that the state of the culture for analysis (steady state vs. non-steady state) could be an important factor for mRNA-protein correlation determination. In line with this are the good

mRNA-protein correlations observed for 56 genes in *E. coli* [4] and at global scale also in *Lactococcus lactis* A-stat cultures [26].

Interestingly, recent time-dependent analysis in yeast cells subjected to osmolarity [149] and oxidative stress [150] revealed good correlations between mRNA and protein expression changes only for up-regulated genes. However, other studies in yeast and bacteria have not detected this kind of behavior [70]. Our results are in accord with the latter as we did not detect differences in mRNA-protein correlation between up- and down-regulated genes. We did not observe different correlations between enzymatic, non-enzymatic and genes with hypothetical protein functions (data not shown).

4.3.2. Functional-genomic responses in central carbon metabolism

Although flux changes reflect the re-organization of metabolism, knowing the preceding gene expression dynamics of both flux catalyzing enzymes and other proteins is important for deciphering the complex cellular regulation patterns enabling the cells to achieve faster growth [13,14].

The major sugar transport system in *E. coli* is the phosphoenolpyruvate: carbohydrate phosphotransferase system (PTS) enabling fast glucose uptake also under glucose-limitation [55,151,152]. Interestingly, our -omics data showed down-regulation of both mRNA and protein of the glucose-specific PTS permease *ptsG* near μ_{max} (Figure 6) while glucose consumption per biomass was maintained constant. This means that the reduced glucose uptake by the PTS system is likely compensated through increasing alternative non-PTS glucose transport. Indeed, we detected up-regulation of alternative glucose transporter galactose permease (*galP*) and predicted transporter *tsgA* (also seen in [6]). However, the former is not a strong candidate as a PTS substitution since expression of glucokinase (*glk*), which is responsible for catalyzing glucose phosphorylation after it is transported into the cell by GalP, was repressed with rising μ. Thus it remains unknown through which mechanism(s) glucose transport is increased at high μ.

In contrast to a change in glucose transport mechanism(s) across the studied range of μ, expression of both mRNAs and proteins in glycolysis were rather insensitive to increasing μ (Figure 6). Constant enzyme expression under strongly increasing flux throughput (5-fold increase of μ in these A-stats) alludes to post-translational regulation of fluxes (see *Control of protein and flux levels to achieve faster growth* for analysis). Expression of PPP components, however, increased with faster growth, most probably to meet the higher NADPH demands (Figure 6). Notably, most of the TCA cycle gene and protein levels were maintained or even increased up to $\mu = 0.40\ h^{-1}$ followed by sudden repression (Figure 6 & Figure 10A), concomitant with achieving maximum q_{CO2} at $\mu = 0.46 \pm 0.02\ h^{-1}$ (Figure 1). This may suggest that there is no limitation at the TCA cycle expression level before $\mu = 0.40\ h^{-1}$ to process the carbon from glycolysis, thus making the hypothesis of TCA cycle bottleneck triggering acetate overflow [81,82] questionable. These observations are in general consistent with the literature [4–7].

Comparison of μ-dependent mRNA and protein expression changes of our A-stat cultures with the only other μ-dependent transcriptome and proteome data set determined in continuous cultures by Ishii *et al.* 2007 [4] showed good correlations of R ~ 0.5 at both levels. Correlating our mRNA expression data with other published μ-dependent transcriptome data sets acquired in continuous cultures showed moderate (R ~ 0.3) and very weak correlations (R ~ 0.1) with [40] and [5], respectively, though, we observed a >70% overlap with [5] in the direction of mRNA expression changes.

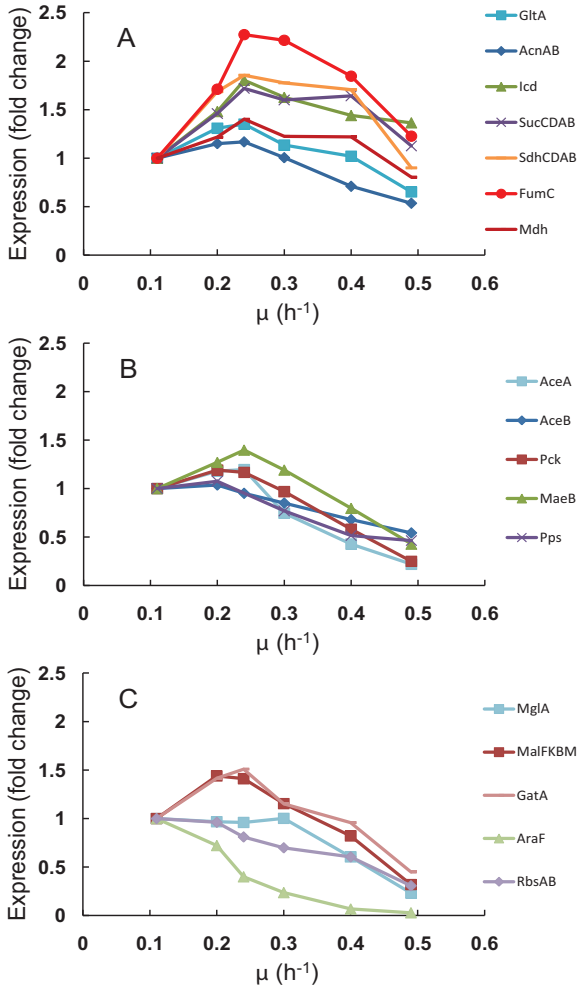

Figure 10. Specific growth rate-dependent TCA cycle, glyoxylate shunt, glyconeogenesis and carbon catabolite repressed protein expression changes in E. coli K-12 MG1655 A-stat cultures. *μ, specific growth rate.* ***A.*** *TCA cycle (average of proteins from the same operon are depicted as one point e.g. AcnAB).* ***B.*** *Glyoxylate shunt (AceA, AceB) and glyconeogenesis.* ***C.*** *Carbon catabolite repressed proteins. Protein data points are average of two independent A-stat experiments, error bars are not shown for better visualization (refer to Additional file 2 in Publication I for standard deviations).*

4.3.3. Activation of carbon catabolite repression

It is well known that the presence of glucose often prevents the use of other, secondary, carbon sources and this preference for glucose is realized through carbon catabolite repression (CCR) [65,66]. More specifically, CCR is defined as a regulatory phenomenon by which the expression of functions for the use of secondary carbon sources and the activities of the corresponding enzymes are reduced in the presence of a preferred carbon source [65,66]. It is plausible that the increasing residual glucose concentration accompanying smooth rise of μ in A-stat could trigger CCR.

Indeed, we observed the beginning of CCR induction prior to the start of acetate overflow. This was indicated by down-regulation (3-fold on average) of CCR-mediated components: alternative (to glucose) substrate transport and utilization systems like galactose (MglAB), maltose (MalBEFKM), galactitol (GatABC), L-arabinose (AraF), D-ribose (RbsAB), C_4-dicarboxylates (DctA) and acetate (ActP, YjcH) (Figure 10C). Moreover, expression of transcription activator Crp (cyclic AMP receptor protein) and Cra (catabolite repressor activator) were reduced 1.5 and 2-time (Figure 11), respectively, similarly to carbon catabolite repressed proteins mentioned above. Simultaneously, components of the gluconeogenesis pathway (Pck, MaeB, Pps) and glyoxylate shunt (GS) enzymes AceA, AceB were repressed with increasing μ (Figure 10B). These results are consistent with transcriptome data of the literature [4–7]. This shows that *E. coli* is exerting CCR leading to the preference of glucose with faster growth even under glucose-limitation.

4.3.4. Expression dynamics of acetate metabolism-related genes

As it is known that acetate overflow is a μ-dependent phenomenon with no excretion of acetate at slow and high excretion at fast growth [4–10,31–34], it is interesting to look into the gene expression dynamics of the main pathways involved in acetate production and consumption around the start of acetate overflow in detail in A-stat cultures.

Since the common theories regarding the cause of acetate overflow assume activation of acetate synthesis pathways [10,31,32,79–82], it was surprising to see in our A-stats a 2-fold down-regulation both on mRNA and protein levels of the two main known acetate synthesis pathways [72]—phosphotransacetylase-acetate kinase (PTA-ACKA) and pyruvate oxidase (POXB)—initiated before the start of acetate overflow (Figure 11). A similar pattern has also been seen in chemostat cultures but without emphasizing the possible physiological consequences [4,5,33]. At the same time, we observed a 10-fold repression of the acetate utilization enzyme acetyl-CoA synthetase (Acs) expression (Figure 11). This substantial 5-fold stronger repression of the acetate consuming Acs in comparison with acetate synthesizing PTA-ACKA suggests disruption of acetate recycling at the PTA-ACS node (highlighted in red in Figure 11). The importance of these observations for the elucidation of the regulation of acetate overflow metabolism will be discussed in the next section.

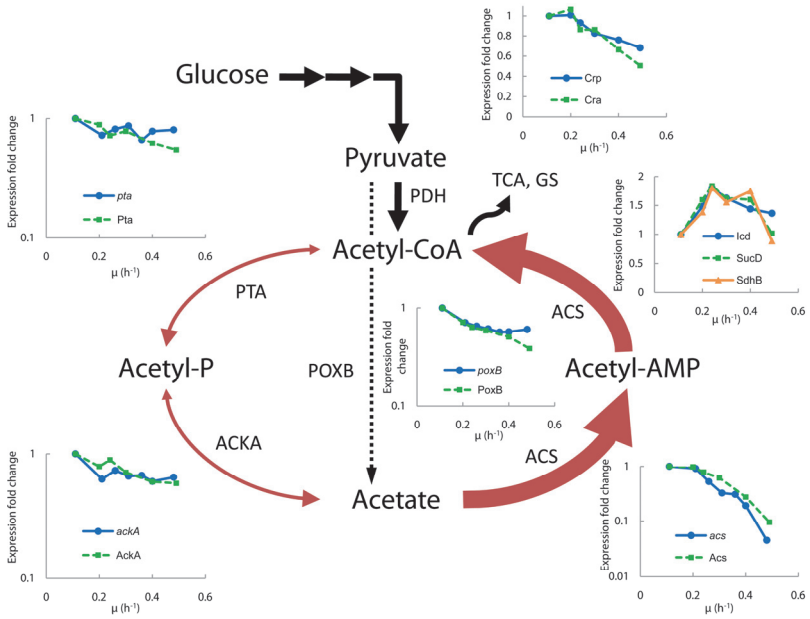

Figure 11. Specific growth rate-dependent acetate synthesis and utilization pathways', selected TCA cycle and carbon catabolite repressed mRNA and protein expression levels in E. coli K-12 MG1655 A-stat experiments. *μ, specific growth rate. PTA, phosphotransacetylase; ACKA, acetate kinase; ACS, acetyl-CoA synthetase; POXB, pyruvate oxidase; PDH, pyruvate dehydrogenase complex; TCA, tricarboxylic acid cycle; GS, glyoxylate shunt; Crp, cyclic AMP receptor protein; Cra, catabolite repressor activator; Icd, isocitrate dehydrogenase; SucD, succinyl-CoA synthethase; SdhB, succinate dehydrogenase. Thickness of red arrows denotes level of ACS and PTA-ACKA pathway repression (thicker line represents stronger repression). Protein data points are average of two independent A-stat experiments, error bars are not shown for better visualization (refer to Additional file 2 in Publication I for standard deviations). mRNA names are in italics.*

4.4. Acetate overflow metabolism (Publications I and III)

Several explanations for acetate overflow metabolism in *E. coli* propose limitations in respiratory capacity [31,32,79], TCA cycle throughput [81,82], energy generation [10,72] or activity of the GS [33,83]. Despite these and other studies over many years, the mechanism and regulation of acetate overflow still remain unclear. Thus one of the more specific aims of this thesis was to gain new insights into the regulation of acetate overflow metabolism using a μ-dependent quantitative multi-omics approach and utilize this knowledge to mitigate carbon wasting into acetate.

4.4.1. Importance of PTA-ACS node

A very important observation in our experiments was the 5-fold stronger repression of the acetate consuming Acs in comparison with acetate synthesizing PTA-ACKA (Figure 11) together with a decline in TCA cycle fluxes (Figure 6) since it suggests

disruption of acetate recycling at the phosphotransacetylase-acetyl-CoA synthetase (PTA-ACS) node. While this node may seem only as a futile cycle (an equivalent amount of ATP to acetate is wasted with acetate recycling), the fact is that numerous metabolic tasks involving the intermediate molecules of this cycle—acetyl phosphate (acetyl-P) and acetyl-AMP (Figure 11)—are essential for *E. coli* growth (Figure 12). For example, both molecules play a critical role in chemotaxis regulation [153–155] and in acetylation of many proteins [156–158], while acetyl-P is additionally vital in pathogenesis [159], protein degradation [160] and can influence the regulation of almost 100 genes in *E. coli* [161]. It is important to note that the only known pathway in *E. coli* for acetyl-P synthesis is the PTA-ACKA [72,155]. Utilizing both acetyl-P and acetyl-AMP in these metabolic tasks through dephosphorylation results in acetate formation. Hence, acetate should be synthesized and consumed simultaneously during growth (not only at fast growth) to maintain proper homeostasis of these metabolites in the PTA-ACS node. Thus we proposed that acetate overflow is triggered by the strong repression of Acs resulting in disruption of the PTA-ACS node due to decreased consumption of acetate by Acs relative to its production by Pta. In fact, a similar hypothesis for acetate overflow has been proposed for *Saccharomyces cerevisiae* [91].

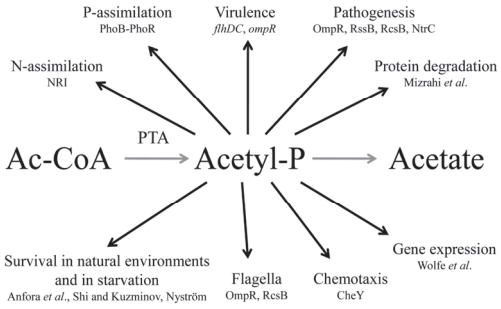

Figure 12. Acetyl phosphate as an important signaling molecule. *Ac-CoA, acetyl-CoA; Acetyl-P, acetyl phosphate; PTA, phosphotransacetylase. Roles of acetyl-P in other processes than stated in the text are also shown – refer to Publication I for detailed explanation.*

4.4.2. Experimental proof of the novel hypothesis for acetate overflow metabolism

We first sought out to experimentally test the latter hypothesis by analyzing μ-dependent acetate consumption capability of *E. coli* using dilution rate stat (D-stat) and two-substrate (glucose + acetic acid) A-stat cultivations. Two substrate A-stat and D-stat experiments directly proved that acetate consumption capability of *E. coli* is μ-dependent as acetate consumption started to decrease from $\mu = 0.25\ h^{-1}$ (Figure S2 in Publication I) and acetate consumption capability rapidly decreased 12-fold around the start of acetate at $\mu = 0.27 \pm 0.02\ h^{-1}$ (Figure 13), respectively. The good correlation between loss of acetate consumption capability and Acs down-regulation is further evidence to the hypothesis of Acs repression causing disruption of PTA-ACS node and triggering acetate overflow (Figure 13).

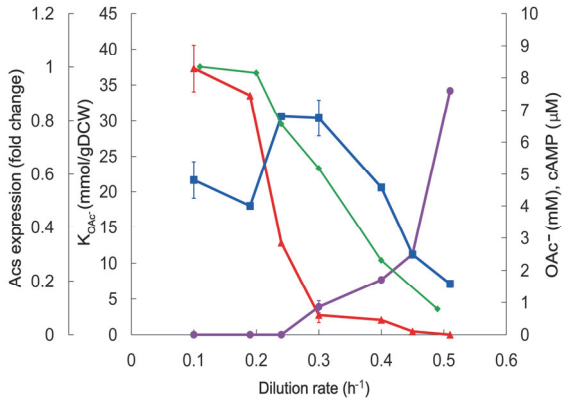

Figure 13. Dilution rate-dependent acetate consumption capability in E. coli K-12 MG1655 D-stat cultures. *gDCW, gram dry cellular weight; K_{OAc-}, acetate consumption per biomass (red triangle); OAc−, acetate concentration in chemostat before the start of acetic acid supplemented medium addition (violet circle); cAMP, cyclic AMP concentration (blue square); Acs, acetyl-CoA synthetase protein expression levels from A-stat (green diamond). Error bars represent standard deviation from two independent D-stat experiments.*

Next, it is reasonable to ask what could cause Acs repression? As acetate is a secondary and less-preferred carbon source compared to glucose [65,66], it is possible that CCR is responsible for the repression of its high affinity scavenging enzyme Acs [93], as also proposed earlier [162]. We saw simultaneous activation of CCR and repression of Acs prior to the start of acetate overflow (Figure 10 & Figure 11). As it is well known that CCR is initiated by the presence of glucose in the environment [65,66], we assumed that the increasing residual glucose concentration accompanying smooth rise of μ in A-stat could trigger Acs down-regulation by CCR. The cAMP-Crp complex is one of the major players in CCR of *E. coli* as cAMP binding to Crp drastically increases its affinity towards activating the promoters of catabolic enzymes, including Acs [65,66,93]. We measured a 1.5-fold decrease in Crp expression with increasing μ (Figure 11) and a strong decline of cAMP levels after the start of acetate overflow close to zero by $\mu = 0.45\ h^{-1}$ (Figure 1) where acetate consumption capability in D-stats was totally lost (Figure 13). As a result of all the previous observations, we proposed a novel hypothesis that acetate overflow is triggered by CCR-mediated Acs repression resulting in decreased consumption of acetate produced by Pta and disruption of the PTA-ACS node.

4.4.3. Coordinated activation of PTA-ACS and TCA cycles strongly reduces acetate overflow

To test if our proposed hypothesis for acetate overflow can also lead mitigation of carbon wasting to acetate, we carried out analysis of several *E. coli* mutant strains. For this, we switched to *E. coli* K-12 BW25113 wild-type (WT), so we would have an appropriate wild-type reference strain for comparing our results obtained with Keio collection single-gene knockout strains [136] and other mutant strains affecting

acetate-related pathways constructed by ourselves. The strains were analyzed in A-stats under identical growth conditions as in the experiments described above (only exception being culture stabilization in chemostat at D = 0.2 h^{-1} in some A-stats) with absolute quantitative exo-metabolome and proteome analyses coupled to MFA. WT and Δ*arcA*, Δ*pka* Δ*arcA* strains were cultivated in four and two independent A-stat experiments, respectively. Other strains were cultivated in single A-stat experiments. We note that here growth characteristics of different strains are compared relative to q_{glc}, instead of μ, since glucose uptake is known to regulate the magnitude of CCR [65,66], as described above, possibly triggering Acs repression and acetate overflow.

We first wanted to verify the hypothesis that *E. coli* actually synthesizes acetate also at low μ but acetate is not excreted because of sufficient consumption by Acs. For this, we analyzed growth of WT with deleted *acs* (Δ*acs*), which should lead to disruption of the PTA-ACS node and loss of acetate recycling. Indeed, the PTA-ACS node disruption in Δ*acs* led to acetate overflow at all q_{glc} values studied, while no acetate accumulated in WT up to q_{glc} ~ 4.2 mmol $gDCW^{-1}$ h^{-1} (Figure 14A). This phenotype of Δ*acs* is consistent with previous observations [33,34] and confirms the hypothesis that acetate is constantly synthesized *in vivo* at all μ. This data estimates that ~0.8 mmol $gDCW^{-1}$ h^{-1} of acetate is constantly recycled in the PTA-ACS node of wild-type *E. coli* to enable rapid regulation of the acetyl-P and acetyl-AMP pools.

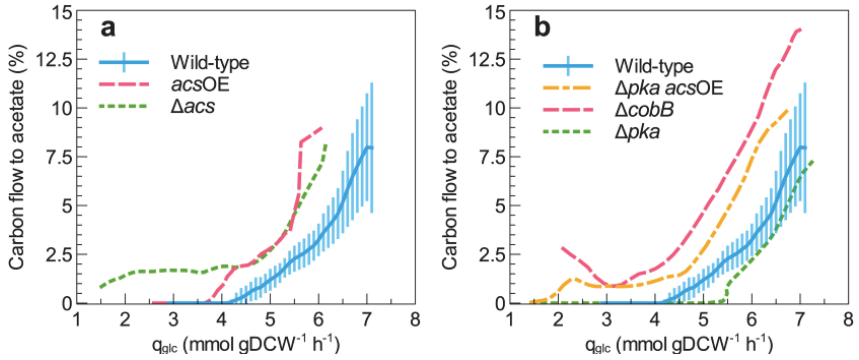

Figure 14. Specific glucose consumption rate-dependent carbon flow to acetate in E. coli K-12 BW25113 wild-type and its Acs-related mutants. *gDCW, gram dry cellular weight; q_{glc}, specific glucose consumption rate. Carbon flow to acetate is percentage from consumed glucose. Lines for each strain are best-fit splines of 9–36 acetate and glucose concentration measurements in each experiment. Error bars represent standard deviation of four independent A-stat experiments. [Publication III] - Reproduced by permission from Springer Science and Business Media.*

As we proposed that acetate overflow is triggered by Acs repression, we next tested the effect of increasing Acs expression on acetate excretion by constructing an *acs* over-expression strain (*acs*OE). However, acetate overflow started at a lower q_{glc} in *acs*OE and the strain also excreted more carbon as acetate than WT at all the q_{glc} studied (Figure 14A). Unfortunately, no clear conclusions could be drawn from literature data with *acs* over-expression strains [33,92]. The lack of a positive effect of

acs over-expression on mitigating acetate overflow could possibly be a result of either the negative effects of too high protein over-expression and/or due to most of the Acs protein pool being inactive. Hence, we next concentrated on the effects regulating Acs protein activity.

Two enzymes in *E. coli* are known to regulate the activity of Acs protein: protein lysine acetyltransferase (Pka) inactivates Acs by acetylation [163], while the NAD^+-dependent regulator protein deacetylase CobB can release Acs from repression by deacetylating it [164]. Thus, to investigate the effects of altered protein acetylation on acetate metabolism, we analyzed growth of Pka (Δ*pka*) and CobB (Δ*cobB*) single-gene knockout mutants of *E. coli*. Δ*cobB* cells, which cannot reactivate Acs by deacetylation, excreted acetate at all q_{glc} studied and more compared to WT (Figure 14B), similarly to Δ*acs*. This further confirms the importance of the state of the Acs pool and an active PTA-ACS node in acetate metabolism. Deletion of Pka should eliminate or at least substantially decrease the fraction of inactivated Acs protein in the cell, possibly leading to higher recycling of acetate in the PTA-ACS node and delayed acetate overflow. Indeed, Δ*pka* showed a postponed start of acetate overflow at q_{glc} = 5.5 compared to ~4.2 mmol $gDCW^{-1}$ h^{-1} in WT (Figure 14B).

Due to the positive effect of *pka* deletion on mitigating acetate metabolism, we next introduced *pka* knockout into the *acs*OE background (Δ*pka acs*OE) to test if the lack of a positive effect of *acs* over-expression on diminishing acetate overflow could have arisen from a substantial part of the Acs pool being inactive in the *acs*OE strain. This did not seem plausible since acetate overflow was not decreased in Δ*pka acs*OE compared to WT (Figure 14B) because this *acs* over-expression strain should possess a fully active Acs pool.

We next hypothesized that acetate overflow could be further reduced in Δ*pka* if Acs levels would be slightly higher, but not too high as in Δ*pka acs*OE. Moreover, even if Δ*pka* cells could recycle more acetate due to higher levels of active Acs, downstream pathways such as TCA cycle could still be limiting in processing the recycled acetyl-CoA in diverting carbon away from acetate. Deletion of the dual transcriptional regulator for anoxic redox control (ArcA) increases both TCA cycle gene expression [5,165] and flux throughput [166]. As *arcA* deletion also leads to about a 2-fold higher expression of *acs* [167], it seemed that deleting *arcA* in the Δ*pka* would hit two birds with one stone by increasing both levels of active Acs and downstream throughput in the TCA cycle. Indeed, the double-knockout strain (Δ*pka* Δ*arcA*) showed even further diminished acetate overflow as the onset of acetate overflow was delayed until q_{glc} ~ 6 mmol $gDCW^{-1}$ h^{-1} compared to q_{glc} ~ 4.2 mmol $gDCW^{-1}$ h^{-1} in WT, and more remarkably, carbon wasting into acetate was more than 4-fold lower compared to WT (2 vs. 8% of total carbon) at maximum q_{glc} (Figure 15A). Absolute quantitative proteome analysis confirmed that Acs expression had increased ~1.7-fold in response to *arcA* deletion (Figure 4 in Publication III). Notably, more CO_2 instead of acetate was produced in Δ*pka* Δ*arcA* (2–8% more than WT) (Figure 15B). Furthermore, MFA showed that higher CO_2 production was the result of increased TCA cycle fluxes (Figure 15C), as expected from an *arcA* deletion [166], while PPP throughput was reduced (Figure 15D).

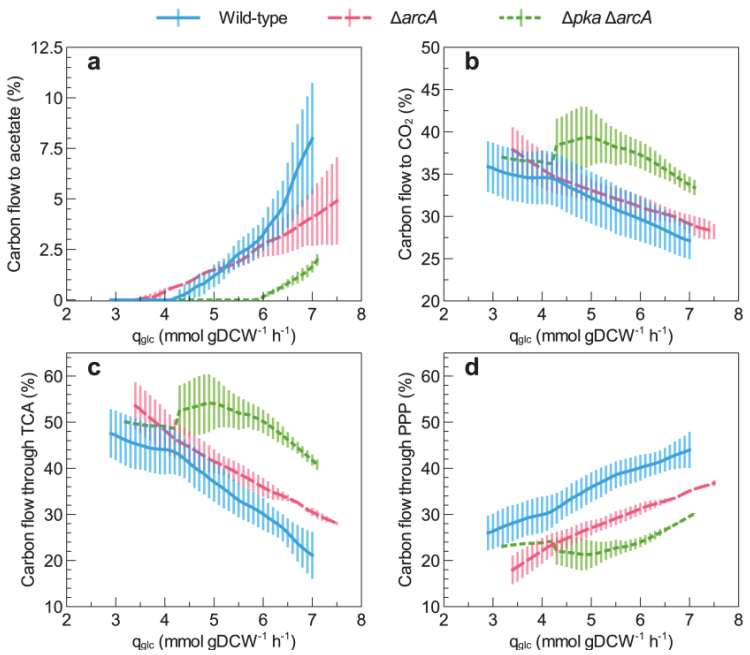

Figure 15. Specific glucose consumption rate-dependent carbon flow through metabolism in E. coli K-12 BW25113 wild-type and its acetate overflow-reduced mutants. *gDCW, gram dry cellular weight; q_{glc}, specific glucose consumption rate; TCA, TCA cycle; PPP, pentose phosphate pathway. The percentage of carbon flow from glucose to acetate (**A**) and CO_2 (**B**). The percentage of carbon flow from glucose through TCA cycle (**C**) and PPP (**D**) represented by suc and gnd fluxes, respectively, expressed as the C-molar percentage of flux through the reaction from consumed carbon. Lines for each strain are best-fit splines of 13–43 glucose and product concentration measurements in each experiment. Error bars represent standard deviation of four and two independent A-stat experiments for wild-type and ΔarcA, Δpka ΔarcA, respectively. [Publication III] - Reproduced by permission from Springer Science and Business Media.*

As acetate overflow could be postponed by a single *arcA* deletion in *E. coli* K-12 MG1655 [5], we also analyzed Δ*arcA* in our BW25113 background to check if the diminished acetate overflow in Δ*pka* Δ*arcA* could actually arise from only the *arcA* deletion. This was not the case as acetate overflow started earlier in Δ*arcA* compared to WT (q_{glc} ~ 3.5 vs. ~ 4.2 mmol gDCW^{-1} h^{-1}) (Figure 15A).

It is important to note that the Δ*pka* Δ*arcA* strain with greatly diminished acetate overflow did not accumulate any other harmful by-product, maintained μ_{max} of WT (Figure 15A) and displayed only ~5% lower Y_{XS} compared to WT. This makes it superior to previous acetate overflow-reduced *E. coli* strains of deletions in main acetate synthesis pathways (*pta* + *ackA* or *poxB*) due to their greatly elevated lactate and formate excretion, reduced μ or Y_{XS} [29,76,87–90]. Furthermore, the Δ*pka* Δ*arcA* strain would permit the production of target molecules in the absence of acetate at higher glucose uptake rates, presumably leading to higher volumetric productivities (~22% higher gDCW L^{-1} h^{-1} compared to WT based on this work).

One of the textbook theories is that acetate overflow in *E. coli* is triggered by limitations in respiratory capacity [31,32,79], meaning that cells start to excrete acetate after reaching maximum q_{O2}. However, our high-resolution A-stat data clearly showed that although there is a change in the slope of q_{O2} with the start of acetate overflow, no saturation of respiratory capacity is observed (Figure 16).

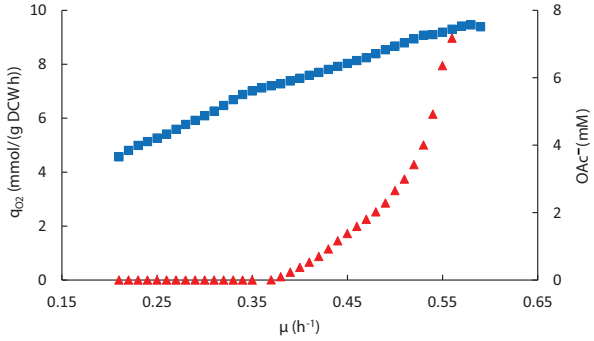

Figure 16. Specific growth rate-dependent respiratory capacity and acetate overflow in E. coli K-12 BW25113 wild-type. *μ, specific growth rate; gDCW, gram dry cellular weight; q_{O2}, specific O_2 consumption rate (blue square); OAc^-, acetate concentration (red triangle).*

It can be concluded from this study of acetate overflow metabolism that a fine-tuned coordination between increasing the recycling capabilities of acetate in the PTA-ACS node through a higher concentration of active acetate-scavenging Acs and downstream metabolism throughput in the TCA cycle leads to greatly reduced acetate overflow in *E. coli*.

4.5. Absolute quantification of transcriptome and proteome (Publication IV)

To have a more comprehensive understanding of cellular processes, a systems biology approach of integrating transcriptome, proteome and fluxome data coupled to models of different levels of detail is needed [14,68,116,117]. This may seem as an immense challenge, but as understanding of both the regulation levels of protein and flux levels and other molecular relationships enabling the cells to modify μ is of instrumental importance towards a more complete description of the control principles of cell metabolism [13,14], more accurate modeling [97,112,113] and successful biotechnology and synthetic biology efforts [17], genome-wide absolute quantitative multi-omics analysis in one study is highly needed. For instance, absolute quantitative proteome (*e.g.* proteins per cell or cell volume) data is required to analyze the allocation principles of cellular proteome resources, the energetic and expression load of certain proteins or pathways, estimate apparent *in vivo* catalytic rates of enzymes and ribosomes etc. What is more, μ-dependent genome-wide absolute quantitative mRNA and proteome data for *E. coli* are missing altogether. Thus we performed μ-dependent absolute quantification of transcriptome and proteome in *E. coli* K-12 MG1655 A-stat cultures.

4.5.1. Global specific growth rate-dependent absolute proteome and transcriptome

Both the intracellular abundance (molecules per cell) and concentration (molecules per fL of biomass) for all the quantified mRNAs and proteins were determined taking into account μ-dependent cell counts and volume, and total RNA and protein fractions in the biomass (Table S1 in Publication IV). Absolute quantification of proteome was performed for 1,185 proteins in chemostat at $\mu = 0.11\ h^{-1}$ using the iBAQ approach [99] and published in [143]. Correlation and coefficient of variation (CV) between two independent A-stat experiments were R = 0.99 and 11%, respectively. Absolute quantification for $\mu = 0.20; 0.30; 0.40; 0.49\ h^{-1}$ was carried out based on relative protein expression data obtained in Publication I which also showed high reproducibility. High confidence of our absolute quantitative proteome data set is indicated by the fact that the sum of all quantified proteins by iBAQ was on average only 10% lower than the cellular total protein concentration determined by Lowry analysis. Genome-wide absolute quantitative transcriptome data for $\mu = 0.11; 0.21; 0.31; 0.40; 0.48\ h^{-1}$ were determined for around 4,300 mRNAs from the DNA microarray data of Publication I. CV between six DNA microarray technical replicates was 11%. All the following data can be found in Table S2 of Publication IV.

Proteomic coverage of protein-coding quantified mRNAs reached 100% with higher mRNA concentrations (Figure 17A) which is consistent with the proposed 'lazy step function'—the ability to detect proteins rising at higher mRNA levels—of protein identification in *E. coli*, *Bacillus subtilis* and higher organisms [70,167]. Both mRNA and protein concentrations spanned approximately three orders of magnitude (Figure 17B) while protein concentrations were on average 1,000-fold higher than their respective mRNAs. A larger dynamic range of protein abundances has been observed in yeast [100] and mammalian cells [70].

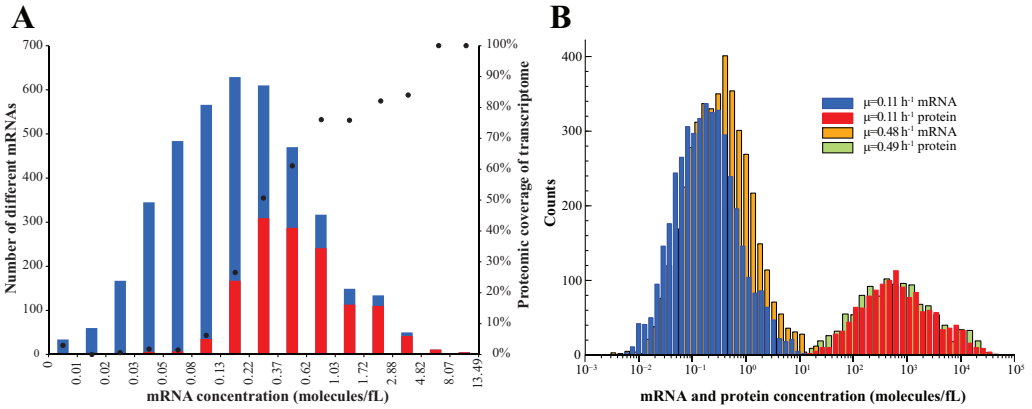

Figure 17. Global absolute proteome and transcriptome characteristics. *μ, specific growth rate.* ***A.*** *Proteomic coverage of protein-coding quantified mRNAs at μ = 0.11 h^{-1}. All quantified protein-coding mRNAs are binned by their concentration. The number of different mRNAs and mRNAs with respectively quantified translation product—protein—belonging to a particular bin is indicated in blue and red, respectively. Black circles denote the proteomic coverage of transcriptome.* ***B.*** *mRNA and protein concentration dynamic range at low and high μ. [Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

As expected with increasing cell size, the sum of mRNAs and proteins per cell increased 4.2- and 2-fold, respectively, from μ = 0.11 to 0.49 h^{-1} (Figure S2A in Publication IV). This is in line with the increase of RNA-to-protein mass ratio with rising μ also observed previously in *E. coli* [97,168]. However, the sum of intracellular mRNA and protein concentrations (molecules per fL) showed a different behavior: 1.8-fold increase for mRNAs and 1.2-fold decrease for proteins was observed (Figure S2B in Publication IV). Similarly, opposite μ-dependence of protein abundances and concentrations has also been observed for unregulated constitutive genes in *E. coli* [114]. The latter indicates a faster increase of cell volume compared to protein abundance with faster growth. Furthermore, the different μ-dependent mRNA and protein abundance and concentration behavior is an important observation for *in silico* modeling approaches where cell metabolism is simulated as a function of μ [112–114].

4.5.2. Correlation of transcriptome and proteome levels

One would expect strong correlation between mRNA and protein levels based on the central dogma of molecular biology, especially for prokaryotes due to simultaneous transcription and translation. So far, correlations of R ~ 0.4–0.7 have been noted in the literature for bacteria, yeasts and multi-cellular organism at population level studies looking either at a small number of genes or lacking direct measurements of both mRNA and protein concentrations (reviewed in [67–70]). Lately, simultaneous quantification of mRNA and protein abundances in *Mycoplasma pneumoniae* [108] and *E. coli* at single cell level [110] revealed correlations of R = 0.52 and 0.77, respectively. Our genome-wide mRNA and protein concentration data show high and

increasing correlations from R = 0.62 to 0.78 in the range of μ studied (Figure 18A), similar to the other *E. coli* continuous culture data set, though for only 56 genes [4]. These values are maximum for our data since correlations were neither improved using nonlinear transformation nor removing outliers (data not shown). Interestingly, mRNA and protein levels correlated better in longer transcription units and for genes located at the 5' end (Figure S18 in Publication IV), similar to *M. pneumoniae* [108].

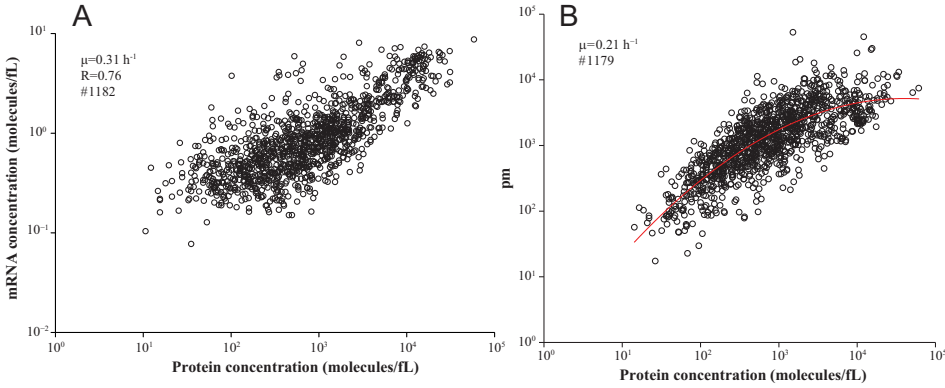

Figure 18. Correlation of protein concentrations with mRNA concentrations and protein-per-mRNA ratios in E. coli K-12 MG1655. *μ, specific growth rate; R, Pearson correlation coefficient; #, number of compared genes; pm, protein-per-mRNA ratio. Red line shows the 3rd order polynomial best fit. Data for all μ in Figure S1 and S3 in Publication IV. [Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

The part of protein abundances which cannot be explained by the correspondence between transcript and protein abundances can be caused by different regulation of mRNA and protein degradation, protein translation, post-translational modification and possible functional requirement for protein binding [67–70]. Of course, experimental and biological noise emerging from the inherent gene expression stochasticity [169] cannot be excluded either [67].

Comparison of absolute proteome data with the literature showed the highest overlap with another *E. coli* steady state study [4] (detailed comparison in [143]). Our transcript abundances correlated highly with other genome-wide *E. coli* calculated values using DNA microarrays in chemostat (R ~ 0.8) [40] and batch cultures (R = 0.91) [102], and modestly (R = 0.52) with recent single cell analysis [110] (data not shown).

4.5.3. Protein-per-mRNA characteristics

The protein-per-mRNA ratio (pm) estimates translation efficiency and changes in pm give indications about the level of gene expression regulation either through protein translation or degradation making it an important molecular characteristic [69]. Values of pm for ~1,200 genes ranged from around 100 to 10,000 at low μ compared to 50 to 4,000 at high μ (Figure 18B) pointing to different levels of post-transcriptional

regulation [67–70] and explain the non-perfect mRNA and protein correlations. Plotting pm values against protein concentrations showed that abundant proteins possess ~100-time higher pm values than low abundant ones regardless of μ with values saturating at around 4,000 and 2,500 at low and high μ, respectively (Figure 18B). A similar difference in pm values between low and high abundant proteins and the saturation effect are seen in yeast [100] and for translation rates in mammalian cells [99]. Median pm decreased 2.3-fold from 1,532 to 656 with μ increasing from 0.11 to 0.48 h^{-1}. The lower ratio determined at $\mu = 0.48\ h^{-1}$ is close to those of log-phase *E. coli* cells [102,110]. The finding of non-constant translation rates is important and already included in recent *in silico* simulations of μ-dependent gene expression [112].

Codon bias *i.e.* non-random occurrence of codons for coding amino acids, is a mechanism for the cell to maximize translation efficiency [67,101] and recently genome-wide correlation between codon adaptation index (CAI) [170], a common estimate for codon bias, and protein expression levels has been shown for *E. coli* [102–106]. However, experiments using synthetic genes show that neither local nor global codon bias have significant effects on mRNA or protein levels [171], although these results have later been objected [172]. Our results are in accord with the previous genome-wide studies as CAI values correlated with protein concentration and pm (Figure S12 in Publication IV), and interestingly increased slightly with rising μ alluding to possible growing pressure for higher translation efficiency for achieving faster growth.

It has been stated that the pm may vary between genes with different function and also change under different conditions [69], as proved by the higher pm ratios for metabolic genes in mammalian cells [173]. We observed that genes belonging to clusters of orthologous groups (COG) [174] of translational machinery (J), energy generation (C) and post-translational modification (O) showed higher pm ratios (Figure S4 in Publication IV) indicating the importance of efficient translation of these enzymes. However, pm values decreased uniformly for all COGs with rising μ (data not shown).

Variable patterns of pm can be affected by protein half-lives known to be influenced by particular amino acids present at the N-terminal end of proteins (the N-end rule) [67]. This has also been reported for *E. coli* with the destabilizing amino acids being Arg, Lys, Leu, Phe, Tyr and Trp [175]. Our analysis of whole protein sequences throughout the range of μ confirms the latter as Arg, Leu, Trp and Glu showed statistically significant enrichment ($p\text{-value} < 0.05$) in proteins with both the lowest pm ratios and concentration (Figure S10 in Publication IV), demonstrating the influence of protein degradation on pm levels.

Our data set did not show significant correlations between pm (translation efficiency) and gene length, protein molecular weight, mRNA levels opposite to what has been proposed previously [69]. Furthermore, weak correlations were found between pm and mRNA half-lives determined in the literature [40,110,176,177] (data not shown).

4.5.4. Characteristics of proteome resource allocation

Knowing that a cell increases expression of protein X 10-fold as a response to an environmental perturbation does not necessarily mean that the expression load of protein X on metabolism is more significant than that of protein Y which is up-regulated only 2-fold. This is because the real metabolic load of protein expression is determined also by the protein amount and that metabolic load would be much higher for protein Y if its abundance is 10-fold higher than that of protein X. Since translational capacity is believed to be rate-limiting for faster growth of *E. coli* [97], it is useful to analyze the allocation principles of proteome resources.

As expected, the most abundant proteins on average over the studied range of μ were involved in translation (elongation factor TufA, ribosomal proteins) and CCM (glycolysis, TCA cycle and PPP) (Table S2 in Publication IV). To find out which proteins are the most and least costly for *E. coli* to express for achieving faster growth, we calculated the protein synthesis cost in ATP (nATP; protein concentration × protein length × 4.306 ATP) for each protein and applied covariance analysis to statistically detect the most and least relevant proteins (see *Materials and methods* for details). The most costly protein for *E. coli* by far with increasing μ was MetE (Figure 19), probably expressed to counteract the detrimental accumulation of homocysteine [78] after the start of acetate overflow leading to increasing acid stress. MetE was followed mainly by ribosomal and amino acid metabolism-related proteins. On the other hand, *E. coli* 'saved' the most energy at higher μ by repression of GS enzymes AceA and AceB, acetate scavenging Acs, enzymes involved in utilization of alternative substrates and several stress response proteins, pointing to the rationale of CCR for regulating proteome resources [178].

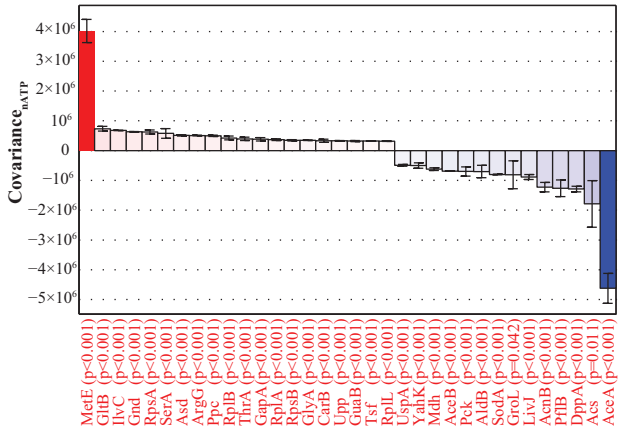

Figure 19. Covariance analysis of protein synthesis cost for the 20 most and least costly proteins for E. coli to achieve faster growth. *nATP, protein synthesis cost. Covariance values are calculated between each protein's nATP values and μ. Red and blue colored bars indicate increasing and decreasing nATP, respectively, with rising μ. Change of nATP is statistically significant for all shown proteins since their covariance values are different from zero. Refer to Figure 21 legend for description of error bars and statistical analysis. [Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

To obtain a functional picture of proteome allocation and expression costs, we grouped all proteins into COG functional classes [174]. Although, translational machinery proteins (J) were by far the most abundant COG group (23–36% of total proteome; $p < 0.0001$), nATP of group J proteins was increasingly exceeded with rising μ by the cost of proteins related to energy generation (C), carbohydrate (G) and amino acid (E) metabolism (Figure 20A; $p < 0.0011$). This is expected since increased substrate utilization, amino acid synthesis and energy generation are needed for faster growth. Turning attention to pathways, proteins of glycolysis, TCA cycle, GS and purine synthesis accounted for both the highest concentration (~19%) and nATP (~27%) of the total proteome (Figure 20B).

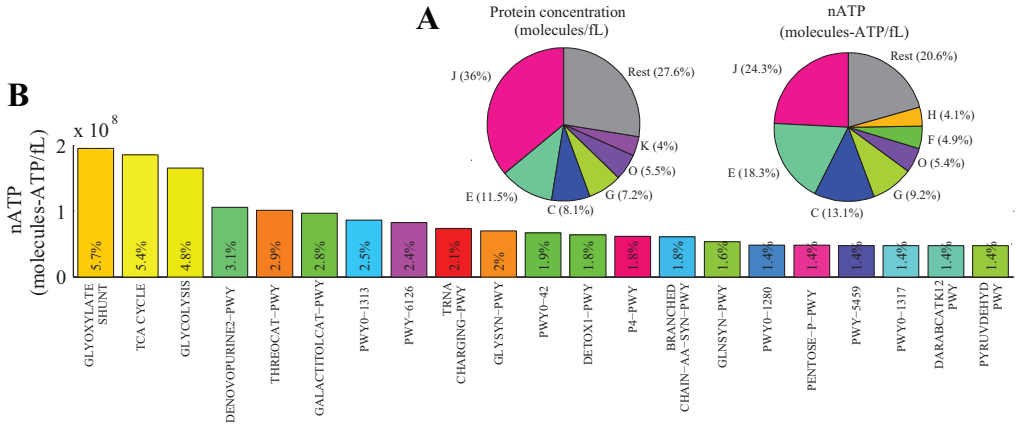

Figure 20. Proteome resource allocation in E. coli K-12 MG1655. *nATP, protein synthesis cost.* ***A.*** *Distribution of protein concentrations and nATP among COG functional classes at $\mu = 0.49\ h^{-1}$. Areas of slices are proportional to absolute values. COG nomenclature: O, post-translational modification; K, transcription; F, nucleotide transport and metabolism; H, coenzyme transport and metabolism; Rest, sum of COGs not shown with independent slices; see text for others.* ***B.*** *Pathways with highest nATP from total proteome synthesis cost at $\mu = 0.11\ h^{-1}$. Percent shows fraction from total cost. See Table S7 in Publication IV for the genes assigned to pathways according to the EcoCyc database* [64]. *Data for all μ in Figure S1 and S3 in Publication IV. [Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

A Pareto principle—top 20% of proteins by abundance accounting for 80% of total protein mass—of proteome resource allocation has been observed in yeast [179], *Leptospira interrogans* [180] and *M. pneumoniae* [108] while the top 20% comprise above 90% in mammalian cells [181,182]. Our data is similar to that of microorganisms, as the top 20% accounted for 76% of protein mass. Interestingly, the 20 most prominent proteins accounted for 19% of protein mass while slightly higher mass fractions of top 20 proteins were reported in all the latterly referred studies (data not shown).

As cells have evolved under energy-limited conditions, the amino acid composition of the proteome should be biased towards containing more amino acids with lower nATP. Indeed, highly expressed proteins in *E. coli* and *B. subtilis* contain more ‘cheap’

(Glu, Asp, Gly) and less 'costly' (Trp, Phe, His, Cys, Leu) amino acids [183]. Our proteome-wide data show the same as the most and least abundant proteins contained more 'cheap' and more 'costly' amino acids, respectively (Figure S11 in Publication IV).

Ribosomes, mainly composed of ribosomal proteins and rRNA, are one of the most important molecules enabling the cells to grow bigger and faster. It has been concluded that ribosomal protein levels are insensitive to μ [184]. However, our data demonstrate 3.4-fold higher median ribosomal protein abundances (7,164 vs. 24,509 molecules per cell) in faster growing cells while accounting for 9–16% of the total protein mass which is close to the estimated value of 21% previously reported for fast growing log-phase *E. coli* cells [55]. Also, our data shows close stoichiometry for ribosomal proteins at all μ values (Figure S7A in Publication IV) and confirms the previously described dimerization of the acidic ribosomal protein RplL [185], also known as RpL7/L12 by being present in roughly double the abundance of other ribosomal proteins at all μ. We also noted 2-fold higher abundance for a 30S subunit ribosomal protein RpsP.

4.5.5. Efficiency of energy generation pathways

Quantification of both μ-dependent absolute proteome and flux values enables to gain unique insights into regulation of energy metabolism by calculating efficiency of energy generation (E_{ATP}) for energy generating pathways. We define E_{ATP} as ATP produced in the pathway per ATP spent for the synthesis of the pathway proteins (molecules-ATP/molecules-ATP) and analyze E_{ATP} dynamics with rising μ for the main energy generating pathways under aerobic growth of *E. coli*: glycolysis, acetate synthesis and the TCA cycle coupled to the respiratory chain (RC) (Table 3). Protein synthesis costs of the ATP producing pathways might be a relevant factor in optimization of the overall strategies of energy generation since ~50% of total ATP for cell proliferation is used for protein synthesis [44,56].

The most efficient energy generating pathway over the studied range of μ was the TCA cycle + RC, for which E_{ATP} varies within 105–152 and is ~5- to 7-fold higher compared to glycolysis (Table 3). In addition, *E. coli* starts to generate additional ATP through acetate overflow after disruption of the PTA-ACS cycle at $\mu = 0.27\ h^{-1}$ and this seems to be beneficial for *E. coli* since E_{ATP} from acetate synthesis exceeds glycolysis at $\mu = 0.4\ h^{-1}$ and even surpasses TCA cycle + RC near μ_{max} (141 vs. 132). This provides a new angle for the potential rationale of acetate overflow for energy generation [10,72] at whole-cell level and is a good example of how new biological knowledge can emerge from integration of genome-wide multi-level quantitative data.

Table 3. Specific growth rate-dependent efficiency of energy production pathways (E_{ATP}) for the main ATP generating pathways under aerobic growth of E. coli K-12 MG1655.

Pathway	μ=0.11 h^{-1}	μ=0.20 h^{-1}	μ=0.30 h^{-1}	μ=0.40 h^{-1}	μ=0.49 h^{-1}
Glycolysis	22.6	21.2	20.3	18.8	19.7
TCA cycle + RC	151.1	116.9	105.5	105.5	132
Acetate synthesis	ND	ND	7.3	41.9	141.2

E_{ATP} is calculated as ATP produced in the pathway per ATP spent for synthesis of the pathway proteins (molecules–ATP/molecules–ATP). ATP produced in the pathway was calculated based on all the ATP producing and consuming fluxes based on MFA. ATP spent for synthesis of the pathway proteins was calculated as a sum of nATPs of all the quantified proteins in the pathway. See Table S7 in Publication IV for the genes assigned to pathways according to the EcoCyc database [64]. *μ, specific growth rate; RC, respiratory chain; ND, not determined since the calculation would be inaccurate due to lack of accurate data for both Pta-AckA and Acs fluxes in the PTA-ACS node before the start of overflow metabolism of acetate at μ = 0.27 h^{-1} while functioning only as an intracellular futile cycle. [Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

4.6. Control of protein and flux levels to achieve faster growth (Publication IV)

Cells have to increase throughput of metabolic fluxes to grow faster. It has become clear, however, that changes in metabolic flux patterns are not a straightforward consequence of transcriptional regulation of enzyme levels [116,122,126,127]. Furthermore, control mechanisms of metabolic fluxes are instrumental for understanding regulation of metabolism since they represent the integrated response of all levels of cellular regulation [41]. Thus one of the aims of this thesis was to determine at which regulation levels of the gene expression cascade (transcriptional (TR), translational (TL), post-TR, post-TL) is control of protein and fluxes levels in different *E. coli* metabolic pathways realized for achieving faster growth.

Our experimental approach was novel since we coupled absolute quantitative global transcriptome and proteome with flux analysis under steady state growth conditions to statistically determine the gene expression regulation levels enabling an organism to grow faster. We used covariance analysis for statistically determining the regulation levels as it describes both the direction and magnitude of mRNA, protein and flux changes with increasing μ making it a suitable statistical method for analysis of absolute quantitative data (see *Materials and methods* for details).

4.6.1. Control of protein concentrations

Impact of TR, post-TR or TL regulation on controlling protein concentrations for *E. coli* to achieve faster growth was quantitatively determined by calculating covariance between each gene's pm (protein-per-mRNA ratio) and μ. The generated covariance values were subjected to statistical hypothesis testing of being significantly (p-value < 0.05) not different or different from zero (all tests were subject to correction by false discovery rate filtering according to the Benjamini–Hochberg procedure [146] at level $\alpha = 0.05$).

A covariance value not different from zero represents TR regulation of protein levels by the pm being constant at all μ. Strikingly, from the total of 1,112 analyzed genes with both quantified mRNA and protein concentrations, only 25% (275) showed TR regulation (Table S3 in Publication IV). Genes involved in GS, NADH metabolism and various degradation pathways showed high fraction of TR regulation (Table S4 in Publication IV). However, for the majority of genes (56%; 627), protein concentrations were controlled at the post-TR level determined by their covariance value being negative and different from zero (Figure 21A), meaning decreasing pm with rising μ (protein concentrations increase less than those of mRNA or decrease more than mRNA). Particularly high enrichment of post-TR regulated genes was observed in COGs of cell cycle (D), translation (J), amino acid metabolism (E) (Table S5 in Publication IV) and amino acid synthesis pathways (Table S4 in Publication IV). Also the protein levels in the pathways carrying the highest flux—glycolysis and TCA cycle—were regulated at post-TR level for >60% of genes. We detected only four genes (0.4%) with TL level regulation indicated by positive covariance values different from zero, meaning increasing pm with rising μ (protein concentrations increase more than those of mRNA or decrease less than mRNA). The small number of proteins which levels were controlled at TL level could be considered as an indication for the faster increase of transcription rate compared to translational rate with growing μ [114].

Similar domination of post-TR regulation of protein concentrations at mRNA and protein levels was lately reported in a study of the fermentative bacterium *L. lactis* [123]. Furthermore, post-TR regulation is the predominant mechanism for controlling pm ratios in *M. pneumoniae* [108], mammalian cells [99] and also for ~37% of genes in *B. subtilis* [111]. Transcription was also shown to have limited control over capacities of key central metabolism enzymes in *E. coli* with faster growth in chemostat cultures [8]. Thus it seems that post-TR regulation of protein levels is significant in organisms with very different growth characteristics, and practically in all the main metabolic subsystems.

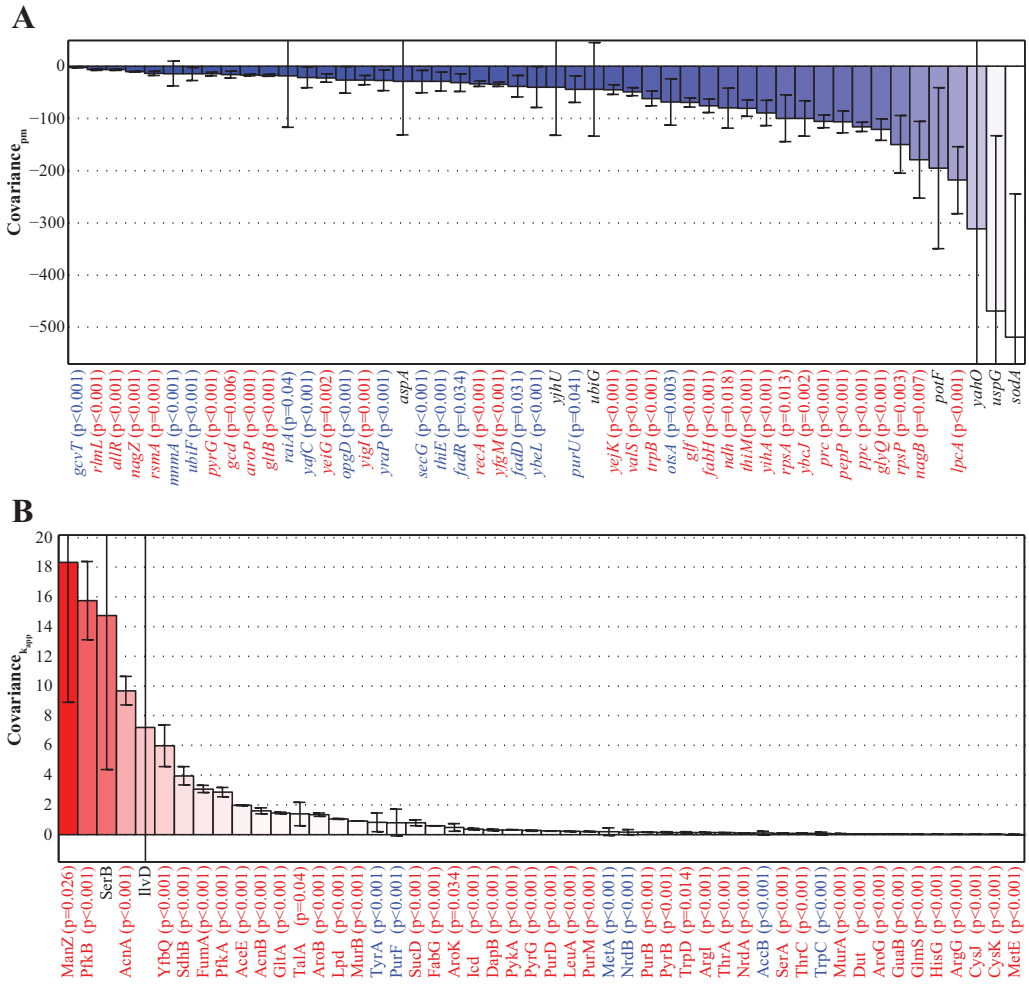

Figure 21. Covariance analysis of protein and flux control levels. *pm, protein-per-mRNA ratio; k_{app}, apparent in vivo catalytic rate of enzyme. Error bars denote 95% confidence intervals of covariance values. p-values represent the results of the statistical hypothesis testing of covariance values being statistically significantly different or non-different from zero.* ***A.*** *Covariance analysis of protein control levels for 50 randomly chosen genes. Covariance values are calculated between each gene's pm values and μ. Red and blue colored bars indicate increasing and decreasing pm with rising μ, respectively. Color of gene name indicates control level: red, post-TR; blue, TR; black, measurement error too large for determination of control level. Data for all genes can be seen in Data S1 in Publication IV.* ***B.*** *Covariance analysis of flux control levels for 50 randomly chosen genes. Covariance values are calculated between each gene's k_{app} values and μ. Red colored bars indicate increasing k_{app} with rising μ. Color of gene name indicates control level: red, post-TL; blue, TL; black, measurement error too large for determination of control level. Data for all genes can be seen in Data S2 in Publication IV. [Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

4.6.2. Control of metabolic flux rates

Next, we moved one layer up to determine the regulation levels of metabolic fluxes for *E. coli* to achieve faster growth. Measurement of both protein concentrations and flux rates in this study enabled to quantitatively determine the impact of change in protein concentration and its catalytic rate for realizing higher flux throughput. For this, we calculated apparent *in vivo* catalytic rates of enzymes (k_{app}, s^{-1}) from the ratio of specific flux rate (mmol/(gDCW h)) and protein concentration (mmol/gDCW). Flux control levels—TL or post-TL—were determined by covariance analysis of changing k_{app} and μ.

Covariance value being statistically not different from zero represents protein control of flux (TL regulation) as protein concentration increases proportionally to the specific flux rate with rising μ (*i.e.* constant k_{app}). Only 9% (18) out of 191 genes under analysis showed flux control through protein levels (TL regulation) (Table S6 in Publication IV) which included genes from glutamine and nucleotide synthesis pathways (Table S4 in Publication IV). Notably, for 10 out of these 18 genes, flux control is at the TR level as both of their pm and k_{app} covariance values were not different from zero. Recently, transcriptional regulation of fluxes was shown in the TCA cycle, GS and acetate excretion but not in PPP using transcription factor mutant strains of *E. coli* [186]. The 10 genes for which we detected TR level control of fluxes did not, however, belong to any of these pathways.

Flux throughput was controlled at the post-TL level for the great majority of genes (81%; 154) shown by positive covariance values different from zero, meaning increasing k_{app} with rising μ (Figure 21B). For nearly all the genes organized into COGs of energy production and conversion (C), translation (J) and high flux pathways—TCA cycle and glycolysis—flux control was achieved through post-TL regulation (Tables S4 and S5 in Publication IV). These are all highly abundant proteins (over 5,000 copies per cell) indicating, similarly to control of protein levels, that *E. coli* has to implement additional regulatory processes for increasing their enzymatic capacities to fulfill the demands of rising biomass and energy synthesis throughput with faster growth. This is also reflected by the observation that the average protein abundance and synthesis cost for post-TL compared to TL-regulated genes were ~2-fold higher (5,190 vs. 2,923 molecules per cell and 6.2×10^6 vs. 3.1×10^6 ATP molecules per cell, respectively). Thus it seems to be energetically favorable for *E. coli* to increase the catalytic capacity of abundant proteins through post-TL processes and save ATP from lower protein synthesis costs. These results are in contrast to an *in silico* study which proposed that low-cost enzymes in *E. coli* are less likely to be post-TR regulated [187].

Predominant post-TL control of several central metabolism fluxes in *E. coli* has been reported using *in vitro* enzyme assays when cells were shifted from low to high D in chemostats [8]. Our results are also in accord with an *E. coli* study based on EcoCyc database [64] information on gene expression regulation, which showed coupling of energy generating reactions to enzymatic regulation, important for short-term maintenance of energy homeostasis [188]. In addition, several studies on yeast using

hierarchical regulation analysis [124] have also concluded that fluxes through glycolytic and fermentative pathways are mainly regulated at the post-TR level [125]. Furthermore, prevalence of post-TL regulation in controlling flux throughput through CCM was also observed in *L. lactis* [123] and has been recently concluded to be probably the primary flux controlling mechanism based on numerous studies [116,126].

Post-TL regulation includes modifications of proteins after translation (post-TL chemical modification or allosteric regulation) or change of its catalytic rate through hyperbolic change of enzyme kinetics solely due to substrate concentration changes [122]. As a substantial amount of quantitative evidence for the importance of post-TL regulation in flux control has accumulated by now, it will be highly relevant in the future to further dissect the specific post-TL regulation mechanisms in action when a reliable non-targeted intracellular metabolome method for *E. coli* continuous cultures, and considerable amount of information on protein specific post-TL modifications and allosteric regulation by reactant metabolites becomes available [116,122,126,127,189].

4.6.3. Apparent *in vivo* catalytic rates of enzymes in central carbon metabolism

Control of fluxes at post-TL level indicates increasing enzymatic capacities through catalytic rates. Indeed, we detected a median 3.7-fold increase of k_{app} of 191 CCM and biosynthetic enzymes (Figure 22) compared to non-changing median protein concentration with rising μ confirming that higher flux throughput for faster growth in *E. coli* is mainly achieved through increased catalytic rates of enzymes.

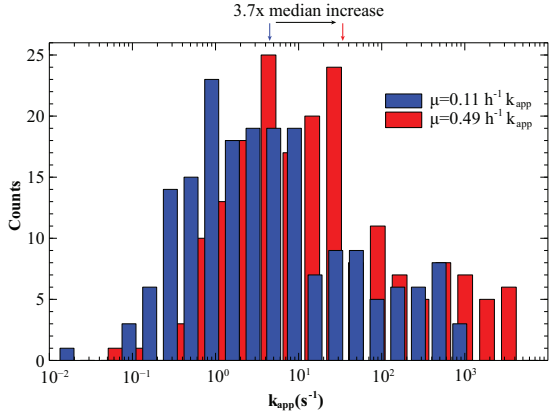

Figure 22. Dynamic range of 191 central carbon metabolism and biosynthetic enzyme apparent in vivo catalytic rates at low and high μ. *μ, specific growth rate; k_{app}, apparent in vivo catalytic rate of enzyme. Blue and red vertical lines with arrowheads above the chart denote the median at μ = 0.11 and 0.49 h^{-1}, respectively. [Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

The same principle for flux control also applies to the most important pathways for aerobic growth of *E. coli*: the average protein concentrations of glycolysis, the TCA cycle, the pentose phosphate pathway (PPP) and acetate synthesis were maintained constant or even decreased with rising μ, and higher flux throughput was clearly

realized through increasing catalytic rates of enzymes (Figure 23). This further supports the conclusion of the relevance of post-TL regulation in metabolic control as these are the pathways carrying the highest flux and responsible for feeding precursor molecules into energy generation and biosynthetic pathways. Our results are in line with a similar recent report for *L. lactis* where a 3.6-fold increase of average k_{app} for CCM and biosynthetic enzymes was detected with 5-time faster growth [123].

When comparing the actual values of k_{app} among the most important pathways, one can see that the average values for the TCA cycle (49–156 s^{-1} from $\mu = 0.11$ to 0.49 h^{-1}, respectively) are higher than for glycolysis (34–110 s^{-1}) and PPP (4–88 s^{-1}) (Figure 23) indicating that pathways with higher contribution of ATP production to the total ATP pool also possess enzymes with higher catalytic rates. Notably, these k_{app} values closely match the range of k_{cat} values for CCM enzymes of *E. coli* measured *in vitro* [190].

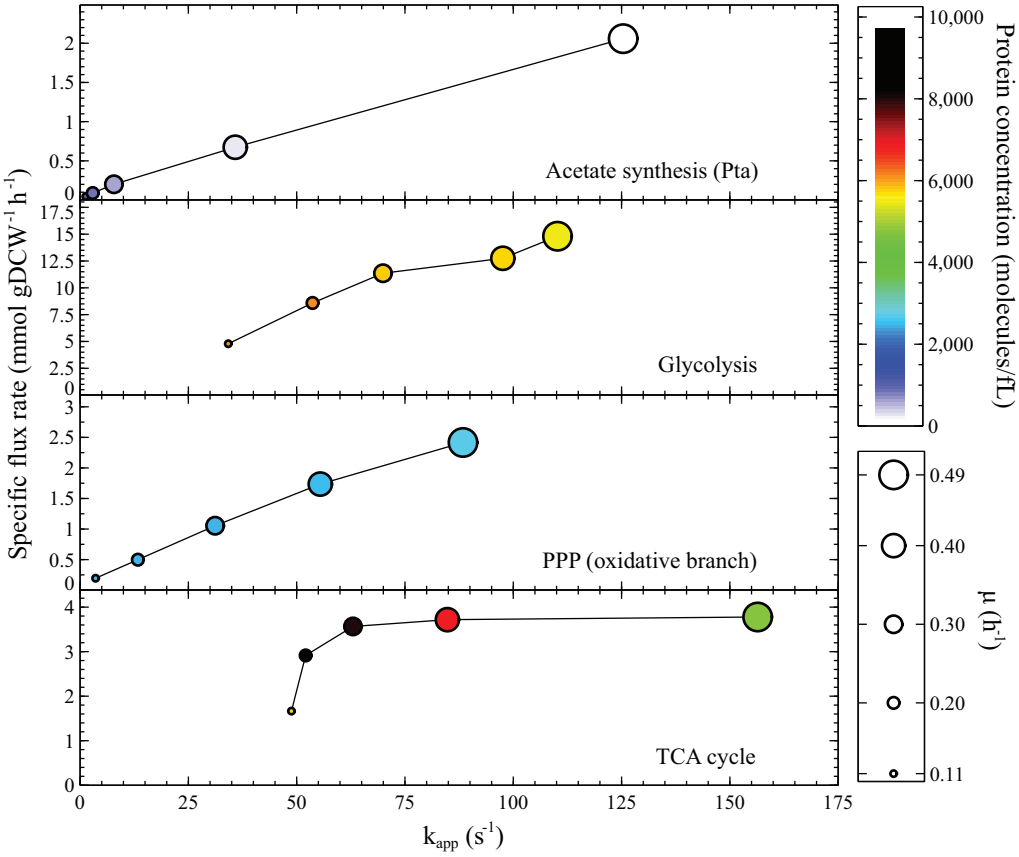

Figure 23. Flux control in the most important pathways for aerobic growth of E. coli. *μ, specific growth rate; gDCW, gram dry cellular weight; k_{app}, apparent in vivo catalytic rate of enzyme. Each circle represents the average value of the pathway if not otherwise noted. See Table S7 in Publication IV for the genes assigned to pathways according to the EcoCyc database* [64]. *[Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

Also at the COG functional class level, k_{app} values for energy metabolism enzymes (C) were an order or two higher than those of biomass monomer synthesis (E, F, H, I) (Figure S6 in Publication IV). Similar higher (~30-fold) catalytic rates for CCM compared to secondary metabolism enzymes is seen when ~2,000 k_{cat} values measured *in vitro* for prokaryotes and eukaryotes were analyzed [190]. Hence, cells maintain higher abundance of proteins required for biomass synthesis (translational machinery, monomer precursor synthesis) whereas enzymatic capacities for energy generating proteins are more likely to be increased through post-TL regulation. Interestingly, we also noted that the pathways carrying the highest flux—glycolysis, TCA cycle and PPP—showed both higher protein concentration and k_{app} compared to biosynthetic pathways (Figure 24). This refers to an evolutionary push towards proteins carrying high flux being more abundant and possessing higher catalytic rates to reduce the cost of protein production, as proposed previously [190,191].

Overall, it seems that under strong nutrient limitation at low μ, metabolism of *E. coli* is on 'standby' mode: protein abundances are high and catalytic rates not saturated for biomass and energy generation so that cells could quickly respond to changing environmental conditions by modifying protein catalytic rates without wasting time for increasing their levels. An analogy for this could be drawn from drag racing: drivers 'pump up' the revolutions of the engines of their cars before the start signal so they could instantly take off with full torque once the light turns green. The same phenomenon is also demonstrated by glucose-pulse experiments where cells are able to rapidly increase μ 3-fold after a substrate pulse [30,192], clearly pointing to the ability to immediately increase catalytic rates if needed.

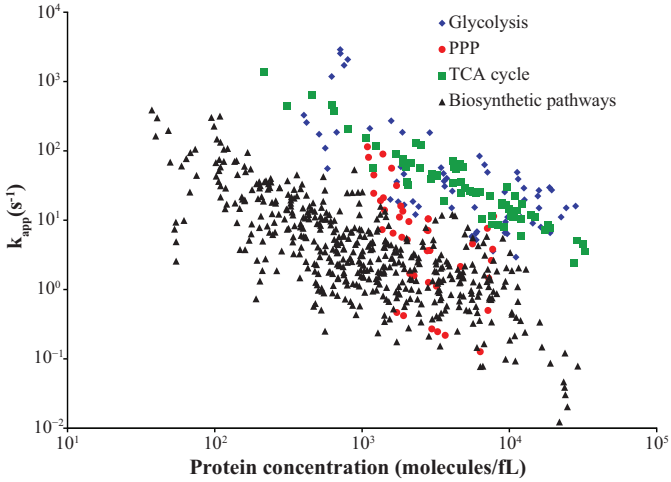

Figure 24. Protein concentrations and their apparent in vivo catalytic rates in pathways carrying the highest flux compared to biosynthetic pathways. *k_{app}, apparent in vivo catalytic rate of enzyme; See Table S7 in Publication IV for the genes assigned to pathways according to the EcoCyc database* [64]. *[Publication IV] - Reproduced by permission of The Royal Society of Chemistry.*

Our quantitative data enable us to investigate patterns of protein catalytic rates within many pathways. In *L. lactis*, first enzymes in central pathways tend to have lower abundances and, hence, higher k_{app} values [123], also seen indirectly in yeast [193]. This hints for possible allosteric control of feedback regulation through the product of the pathway to precisely control the flux [187]. Our data are in accordance with the latter and provide further pathway-specific observations. Both initial enzymes of glycolysis and PPP, Pgi and Zwf, respectively, possess higher k_{app} values compared to subsequent enzymes probably for enabling strict distribution of carbon flow between these pathways for fast switching between the need for more ATP (glycolysis) or NADPH (PPP). One would expect that a biosynthesis pathway is activated when its product is needed and flux is controlled by its first enzyme. However, a difference among biosynthetic pathways was observed: longer pathways (*e.g.* purine, aromatic amino acids, Arg, Lys) showed higher k_{app} values for their first enzymes which was not seen for shorter ones (*e.g.* Ser, Thr, Leu, His). It might be useful for the cell to control flux through initial enzymes in longer pathways to realize a fast response as opposed to a time-consuming situation where all the enzyme levels are maintained low and increased uniformly once the pathway capacity has to be enhanced. More detailed analysis could be found in Publication IV.

4.6.4. Ribosomal translation rate

Faster growing cells are bigger since they need more 'catalytic units' to process nutrients and synthesize biomass faster. For this, a cell needs to increase its translational capacity with rising μ for maintaining the necessary concentrations of enzymes, the cellular 'workhorses'. How is this achieved?

In our experiments, the median ribosomal protein concentration increased only 1.5-fold with 5-time faster growth (Figure S7A in Publication IV) indicating the need for higher translation rates to maintain sufficient translational capacity. Indeed, estimating ribosomal translation rates by calculating k_{app} for ribosomal proteins revealed 2.8-time (median; CV = 11% among 52 quantified ribosomal proteins) higher values with increasing μ (Figure S7B in Publication IV), demonstrating that increasing the concentration of ribosomal proteins is not sufficient for achieving faster growth and, therefore, also translation rates have to be increased. The observation of increased translation rates in faster growing cells is in concordance with the literature using indirectly calculated values [55,168,194].

5. CONCLUSIONS

The main conclusion of this thesis is that the developed comprehensive systems biology approach of absolute quantitative multi-omics analyses coupled to advanced continuous cultivation and computational methods was validated in elucidation of novel regulation principles of cell metabolism which allows further successful metabolic engineering of strains with superior characteristics at a new level.

Other more specific conclusions resulting from this thesis are:

I It was confirmed that highly reproducible accelerostat (A-stat) cultures precisely detect metabolic switch-points (*e.g.* start of acetate overflow), elucidate the dynamics of metabolism and describe steady state physiology equally to chemostat, making A-stat suitable for steady state physiology studies and screening bioprocesses.

II Detailed carbon balance analysis and experimentally determined μ-dependent biomass composition increased accuracy of metabolic flux analysis and enabled to propose that despite increasing carbon wasting with rising μ, *E. coli* could maintain a constant biomass yield through simultaneous reduction of non-growth associated ATP production.

III We propose a novel hypothesis that acetate overflow metabolism in *E. coli* is triggered by carbon catabolite repression-mediated acetyl-CoA synthetase (Acs) repression resulting in decreased consumption of acetate produced by Pta and disruption of the phosphotransacetylase-acetyl-CoA synthetase (PTA-ACS) node.

IV Fine-tuned coordination between increasing the recycling capabilities of acetate in the PTA-ACS node through a higher concentration of active acetate-scavenging Acs and downstream metabolism throughput in the TCA cycle strongly reduces acetate overflow in *E. coli*.

V *E. coli* K-12 BW25113 Δ*pka* Δ*arcA* strain with increased acetate recycling and TCA cycle throughput is superior to previous acetate overflow-reduced *E. coli* strains since it does not accumulate any other detrimental by-product and maintains μ_{max} of wild-type presumably leading to higher volumetric productivities in bioprocesses in the absence of acetate.

VI High correlations (R ~ 0.8) between both genome-wide mRNA and protein expression changes with rising μ and concentrations observed under strictly defined and controlled growth conditions suggest that the state of the culture could be an important factor in mRNA-protein correlation analysis. We also conclude that *E. coli* achieves faster growth through increasing catalytic and translation rates of proteins by predominantly controlling protein abundances and

metabolic flux rates at post-transcriptional and post-translational levels, respectively. This serves the basis for the development of more sophisticated whole-cell models through including the complex regulation layers of gene expression and metabolic fluxes dependent on the physiological state of the cell.

VII *E. coli* enzymes carrying high fluxes seem to be more abundant and also possess higher apparent *in vivo* catalytic rates, suggesting an evolutionary push to reduce protein synthesis costs.

ACKNOWLEDGEMENTS

In parallel with my Ph.D. coursework in Tallinn University of Technology, all the research presented in this thesis was carried out at the Competence Center of Food and Fermentation Technologies (CCFFT). I was lucky enough to be invited to join this company already during my second year of Bachelor studies by Prof. Raivo Vilu. I am thankful for all the people who have made CCCFT happen and prosper for so many years which has enabled me to do cutting-edge research there. I am most grateful to Raivo for making this possible. Moreover, I owe him many thanks for triggering my interest towards science by his very passionate and catching teaching, and I feel very lucky to have had such an open-minded and supportive supervisor. I thank Raivo for giving me freedom in research which has built valuable independent skills in me. Raivo is a true visionary leader from whom you can learn infinitely. Importantly, we have also had some very good times together outside the lab.

I am also very grateful to my co-supervisor Kaarel Adamberg for the great journey so far. Kaarel taught me all the basics of fermentation and has always kept me down-to-earth while supporting some risky ideas. I thank Kaarel for his persistent supervising and for many long discussions on science in general or on some specific aspects.

My one-year internship in Genomatica was a priceless experience and partly I owe this to my supervisors there Steve Van Dien and Gian Oddone. They taught me a lot about metabolic engineering and fermentation in the industry.

I have very much enjoyed working together with my colleagues in our research group in CCFFT. I am most grateful to Petri-Jaan, together with whom I enrolled to the university. I probably would not have come that far without having such a great colleague and friend. His passion and professional attitude has been very inspiring. Many thanks go to Ranno for our very enjoyable joint research. I am also very thankful to Karl for teaching me a lot while supervising him, and the same goes to other students I have had the honor to mentor. Lots of cheers go to my other (ex-) colleagues Liisa, Kadri, Indrek, Andrus, Sten, Janar, Kristo, Andres, Merli, Dagmar and Gethe from whom I have learned a lot. My thanks also go to all my former and current colleagues in CCFFT for all the fun experiences in and outside the lab.

Finally, I express my greatest gratitude to my parents and my sister, and other family for their encouragement and support throughout my studies. I am very proud to be one of our family. I thank my wife Annika for her love and for understanding that the effort of becoming a scientists is a 24/7 thing. Of course, I also thank all my dear friends in Estonia and abroad who have made my life a lot more fun and exciting. I hope I have given you all something in return.

The financial support for this work was provided by the European Regional Development Fund project EU29994; Ministry of Education, Estonia, through the grant SF0140090s08; Estonian Science Foundation through grants G8165 and G9192. These studies were partially supported by European Social Fund's Doctoral Studies, Internationalisation Programme DoRa and Graduate School "Functional materials and technologies" receiving funding from the European Social Fund under project 1.2.0401.09-0079 in University of Tartu, Estonia.

BIBLIOGPRAHY

1 Mongold JA & Lenski RE (1996) Experimental Rejection of a Nonadaptive Explanation for Increased Cell Size in *Escherichia coli*. *J. Bacteriol.* **178**, 5333–5334.

2 Volkmer B & Heinemann M (2011) Condition-Dependent Cell Volume and Concentration of *Escherichia coli* to Facilitate Data Conversion for Systems Biology Modeling. *PLoS One* **6**, e23126.

3 Paalme T, Kahru A, Elken R, Vanatalu K, Tiisma K & Vilu R (1995) The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate (A-stat). *J. Microbiol. Methods* **24**, 145–153.

4 Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, Ho PY, 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 (2007) Multiple High-Throughput Analyses Monitor the Response of *E. coli* to Perturbations. *Science.* **316**, 593–597.

5 Vemuri GN, Altman E, Sangurdekar DP, Khodursky AB & Eiteman MA (2006) Overflow metabolism in *Escherichia coli* during steady-state growth: transcriptional regulation and effect of the redox ratio. *Appl. Environ. Microbiol.* **72**, 3653–3661.

6 Nahku R, Valgepea K, Lahtvee P-J, Erm S, Abner K, Adamberg K & Vilu R (2010) Specific growth rate dependent transcriptome profiling of *Escherichia coli* K12 MG1655 in accelerostat cultures. *J. Biotechnol.* **145**, 60–65.

7 Yao R, Hirose Y, Sarkar D, Nakahigashi K, Ye Q & Shimizu K (2011) Catabolic regulation analysis of *Escherichia coli* and its *crp*, *mlc*, *mgsA*, *pgi* and *ptsG* mutants. *Microb. Cell Fact.* **10**, 67.

8 Vazquez A, Beg QK, Demenezes MA, Bar-Joseph Z, Ernst J, Barabási A-L, Boros LG & Oltvai ZN (2008) Impact of the solvent capacity constraint on *E. coli* metabolism. *BMC Syst. Biol.* **2**, 7.

9 Nanchen A, Sauer U & Schicker A (2006) Nonlinear Dependency of Intracellular Fluxes on Growth Rate in Miniaturized Continuous Cultures of *Escherichia coli*. *Appl. Environ. Microbiol.* **72**, 1164–1172.

10 Kayser A, Weber J, Hecht V & Rinas U (2005) Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. I. Growth-rate-dependent metabolic efficiency at steady state. *Microbiology* **151**, 693–706.

11 Hoskisson PA & Hobbs G (2005) Continuous culture--making a comeback? *Microbiology* **151**, 3153–3159.

12 Bull AT (2010) The renaissance of continuous culture in the post-genomics age. *J. Ind. Microbiol. Biotechnol.* **37**, 993–1021.

13 Phillips R & Milo R (2009) A feeling for the numbers in biology. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21465–21471.

14 Neidhardt FC (1999) Bacterial Growth: Constant Obsession with dN/dt. *J. Bacteriol.* **181**, 7405–7408.

15 Eiteman MA & Altman E (2006) Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol.* **24**, 530–536.

16 Clomburg JM & Gonzalez R (2010) Biofuel production in *Escherichia coli*: the role of metabolic engineering and synthetic biology. *Appl. Microbiol. Biotechnol.* **86**, 419–434.

17 Scott M & Hwa T (2011) Bacterial growth laws and their applications. *Curr. Opin. Biotechnol.* **22**, 559–565.

18 Monod J (1950) La technique de culture continue, théorie et applications. *Ann. Inst. Pasteur*. **79**, 390–410.

19 Novick A & Szilard L (1950) Description of the chemostat. *Science* **112**, 715–716.

20 Van der Sluis C, Westerink BH, Dijkstal MM, Castelein SJ, van Boxtel AJ, Giuseppin ML, Tramper J & Wijffels RH (2001) Estimation of steady-state culture characteristics during acceleration-stats with yeasts. *Biotechnol. Bioeng.* **75**, 267–275.

21 Albergaria H, Torrao A, Hogg T & Girio F (2003) Physiological behaviour of *Hanseniaspora guilliermondii* in aerobic glucose-limited continuous cultures. *FEMS Yeast Res.* **3**, 211–216.

22 Barbosa MJ, Hoogakker J & Wijffels RH (2003) Optimisation of cultivation parameters in photobioreactors for microalgae cultivation using the A-stat technique. *Biomol. Eng.* **20**, 115–123.

23 Barbosa MJ, Zijffers JW, Nisworo A, Vaes W, Wijffels RH & van Schoonhoven J (2005) Optimization of biomass, vitamins, and carotenoid yield on light energy in a flat-panel reactor using the A-stat technique. *Biotechnol. Bioeng.* **89**, 233–242.

24 Girbal L, Rols JL & Lindley ND (2000) Growth rate influences reductive biodegradation of the organophosphorus pesticide demeton by *Corynebacterium glutamicum*. *Biodegradation* **11**, 371–376.

25 Nahku R (2012) Validation of critical factors for the quantitative characterization of bacterial physiology in accelerostat cultures. *PhD thesis, Tallinn University of Technology*.

26 Lahtvee P-J, Adamberg K, Arike L, Nahku R, Aller K & Vilu R (2011) Multi-omics approach to study the growth efficiency and amino acid metabolism in *Lactococcus lactis* at various specific growth rates. *Microb. Cell Fact.* **10**, 12.

27 Emmerling M, Dauner M, Ponti A, Hochuli M, Fiaux J, Szyperski T, Wüthrich K, Bailey JE & Sauer U (2002) Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. *J. Bacteriol.* **184**, 152–164.

28 Yang C, Hua Q, Baba T, Mori H & Shimizu K (2003) Analysis of *Escherichia coli* anaplerotic metabolism and its regulation mechanisms from the metabolic responses to altered dilution rates and phosphoenolpyruvate carboxykinase knockout. *Biotechnol. Bioeng.* **84**, 129–144.

29 Abdel-Hamid AM, Attwood MM & Guest JR (2001) Pyruvate oxidase contributes to the aerobic growth efficiency of *Escherichia coli*. *Microbiology* **147**, 1483–1498.

30 Sunya S, Delvigne F, Uribelarrea J-L, Molina-Jouve C & Gorret N (2012) Comparison of the transient responses of *Escherichia coli* to a glucose pulse of various intensities. *Appl. Microbiol. Biotechnol.* **95**, 1021–1034.

31 Paalme T, Elken R, Kahru A, Vanatalu K & Vilu R (1997) The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach. *Antonie Van Leeuwenhoek* **71**, 217–230.

32 Varma A & Palsson BØ (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **60**, 3724–3731.

33 Shin S, Chang D & Pan JG (2009) Acetate Consumption Activity Directly Determines the Level of Acetate Accumulation During *Escherichia coli* W3110 Growth. *J. Microbiol. Biotechnol.* **19**, 1127–1134.

34 Renilla S, Bernal V, Fuhrer T, Castaño-Cerezo S, Pastor JM, Iborra JL, Sauer U & Cánovas M (2011) Acetate scavenging activity in *Escherichia coli*: interplay of acetyl-CoA synthetase and the PEP-glyoxylate cycle in chemostat cultures. *Appl. Microbiol. Biotechnol.* **93**, 2109–2124.

35 Chen R, Yap WM, Postma PW & Bailey JE (1997) Comparative studies of *Escherichia coli* strains using different glucose uptake systems: Metabolism and energetics. *Biotechnol. Bioeng.* **56**, 583–590.

36 Yates RA & Pardee AB (1956) Control of pyrimidine biosynthesis in *Escherichia coli* by a feed-back mechanism. *J. Biol. Chem.* **221**, 757–770.

37 Womack JE & O'Donovan GA (1978) Orotic acid excretion in some wild-type strains of *Escherichia coli* K-12. *J. Bacteriol.* **136**, 825–827.

38 Cheng K-K, Lee B-S, Masuda T, Ito T, Ikeda K, Hirayama A, Deng L, Dong J, Shimizu K, Soga T, Tomita M, Palsson BØ & Robert M (2014) Global metabolic network reorganization by adaptive mutations allows fast growth of *Escherichia coli* on glycerol. *Nat. Commun.* **5**, 3233.

39 Taymaz-Nikerel H, Borujeni AE, Verheijen PJT, Heijnen JJ & van Gulik WM (2010) Genome-derived minimal metabolic models for *Escherichia coli* MG1655 with estimated in-vivo respiratory ATP stoichiometry. *Biotechnol. Bioeng.* **107**, 369–381.

40 Esquerré T, Laguerre S, Turlan C, Carpousis AJ, Girbal L & Cocaign-Bousquet M (2013) Dual role of transcription and transcript stability in the regulation of gene expression in *Escherichia coli* cells cultured on glucose at different growth rates. *Nucleic Acids Res.* **42**, 2460–2472.

41 Sauer U (2006) Metabolic networks in motion: ^{13}C-based flux analysis. *Mol. Syst. Biol.* **2**, 62.

42 Long CP & Antoniewicz MR (2014) Metabolic flux analysis of *Escherichia coli* knockouts: lessons from the Keio collection and future outlook. *Curr. Opin. Biotechnol.* **28**, 127–133.

43 Sauer U, Canonaco F, Heri S, Perrenoud A & Fischer E (2004) The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. *J. Biol. Chem.* **279**, 6613–6619.

44 Russell JB & Cook GM (1995) Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol. Rev.* **59**, 48–62.

45 Russell JB (2007) The Energy Spilling Reactions of Bacteria and Other Organisms. *J. Mol. Microbiol. Biotechnol.* **13**, 1–11.

46 Chao YP & Liao JC (1994) Metabolic responses to substrate futile cycling in *Escherichia coli*. *J. Biol. Chem.* **269**, 5122–5126.

47 Torres J, Guixe V & Babul J (1997) A mutant phosphofructokinase produces a futile cycle during gluconeogenesis in *Escherichia coli*. *Biochem. J.* **684**, 675–684.

48 Sigala JC, Flores S, Flores N, Aguilar C, de Anda R, Gosset G & Bolívar F (2009) Acetate metabolism in *Escherichia coli* strains lacking phosphoenol-pyruvate:carbohydrate phosphotransferase system; evidence of carbon recycling strategies and futile cycles. *J. Mol. Microbiol. Biotechnol.* **16**, 224–235.

49 Tomashek JJ & Brusilow WS (2000) Stoichiometry of energy coupling by proton-translocating ATPases: a history of variability. *J. Bioenerg. Biomembr.* **32**, 493–500.

50 Ferguson SJ (2010) ATP synthase: from sequence to ring size to the P/O ratio. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 16755–16756.

51 Špitsmeister M, Adamberg K & Vilu R (2010) UPLC/MS based method for quantitative determination of fatty acid composition in Gram-negative and Gram-positive bacteria. *J. Microbiol. Methods* **82**, 288–295.

52 Pramanik J & Keasling JD (1997) Stoichiometric model of *Escherichia coli* metabolism: incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnol. Bioeng.* **56**, 398–421.

53 Pramanik J & Keasling JD (1998) Effect of *Escherichia coli* biomass composition on central metabolic fluxes predicted by a stoichiometric model. *Biotechnol. Bioeng.* **60**, 230–238.

54 Zhao J & Shimizu K (2003) Metabolic flux analysis of *Escherichia coli* K12 grown on ^{13}C-labeled acetate and glucose using GC-MS and powerful flux calculation method. *J. Biotechnol.* **101**, 101–117.

55 Bremer H & Dennis P (1996) Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (Neidhardt F, Curtiss RI, Ingraham J, Lin E, Low K, Magasanik B, Reznikoff W, Riley M, Schaechter M, & Umbarger H, eds), 2nd ed., pp. 1553–1569. ASM Press, Washington, DC.

56 Cox RA (2004) Quantitative relationships for specific growth rates and macromolecular compositions of *Mycobacterium tuberculosis*, *Streptomyces coelicolor* A3(2) and *Escherichia coli* B/r: an integrative theoretical approach. *Microbiology* **150**, 1413–1426.

57 Fuhrer T & Sauer U (2009) Different Biochemical Mechanisms Ensure Network-Wide Balancing of Reducing Equivalents in Microbial Metabolism. *J. Bacteriol.* **191**, 2112–2121.

58 Stephanopoulos G, Aristidou A & Nielsen J (1998) Review of Cellular Metabolism. In *Metabolic Engineering: Principles and Methodologies*. Academic Press, San Diego, USA.

59 Stouthamer AH (1973) A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie Van Leeuwenhoek* **39**, 545–565.

60 Van Bodegom P (2007) Microbial maintenance: a critical review on its quantification. *Microb. Ecol.* **53**, 513–523.

61 Wodke JAH, Puchałka J, Lluch-Senar M, Marcos J, Yus E, Godinho M, Gutiérrez-Gallego R, Martins dos Santos VAP, Serrano L, Klipp E & Maier T (2013) Dissecting the energy metabolism in *Mycoplasma pneumoniae* through genome-scale metabolic modeling. *Mol. Syst. Biol.* **9**, 653.

62 Hempfling W & Mainzer S (1975) Effects of varying the carbon source limiting growth on yield and maintenance characteristics of *Escherichia coli* in continuous culture. *J. Bacteriol.* **123**, 1076–1087.

63 Andersen K & von Meyenburg K (1980) Are growth rates of *Escherichia coli* in batch cultures limited by respiration? *J. Bacteriol.* **144**, 114–123.

64 Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martínez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Schröder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus RP, Paulsen I & Karp PD (2013) EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* **41**, D605–12.

65 Görke B & Stülke J (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol* **6**, 613–624.

66 Narang A (2009) Quantitative effect and regulatory function of cyclic adenosine 5′-phosphate in *Escherichia coli*. *J. Biosci.* **34**, 445–463.

67 Maier T, Güell M & Serrano L (2009) Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* **583**, 3966–3973.

68 Zhang W, Li F & Nie L (2010) Integrating multiple "omics" analysis for microbial biology: application and methodologies. *Microbiology* **156**, 287–301.

69 De Sousa Abreu R, Penalva LO, Marcotte EM & Vogel C (2009) Global signatures of protein and mRNA expression levels. *Mol. Biosyst.* **5**, 1512–1526.

70 Vogel C & Marcotte EM (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227–232.

71 Diez-Gonzalez F & Russell JB (1997) Effects of carbonylcyanide-m-chlorophenylhydrazone (CCCP) and acetate on *Escherichia coli* O157:H7 and K-12: uncoupling versus anion accumulation. *FEMS Microbiol. Lett.* **151**, 71–76.

72 Wolfe AJ (2005) The acetate switch. *Microbiol. Mol. Biol. Rev.* **69**, 12–50.

73 Luli GW & Strohl WR (1990) Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Appl. Environ. Microbiol.* **56**, 1004–1011.

74 Nakano K, Rischke M, Sato S & Märkl H (1997) Influence of acetic acid on the growth of *Escherichia coli* K12 during high-cell-density cultivation in a dialysis reactor. *Appl. Microbiol. Biotechnol.* **48**, 597–601.

75 Koh BT, Nakashimada U, Pfeiffer M & Yap MGS (1992) Comparison of acetate inhibition on growth of host and recombinant *E. coli* K12 strains. *Biotechnol. Lett*. **14**, 1115–1118.

76 Contiero J, Beatty CM, Kumari S, DeSanti C, Strohl WR & Wolfe AJ (2000) Effects of mutations in acetate metabolism in high-cell-density growth of *Escherichia coli*. *J. Ind. Microbiol. Biotechnol.* **24**, 421–430.

77 March JC, Eiteman MA & Altman E (2002) Expression of an anaplerotic enzyme, pyruvate carboxylase, improves recombinant protein production in *Escherichia coli*. *Appl. Environ. Microbiol.* **68**, 5620–5624.

78 Roe AJ, O'Byrne CP, McLaggan D & Booth IR (2002) Inhibition of *Escherichia coli* growth by acetic acid: a problem with methionine biosynthesis and homocysteine toxicity. *Microbiology* **148**, 2215–2222.

79 Han K, Lim HC & Hong J (1992) Acetic acid formation in *Escherichia coli* fermentation. *Biotechnol. Bioeng.* **39**, 663–671.

80 Farmer WR & Liao JC (1997) Reduction of aerobic acetate production by *Escherichia coli*. *Appl. Environ. Microbiol.* **63**, 3205–3210.

81 Majewski RA & Domach MM (1990) Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotechnol. Bioeng.* **35**, 732–738.

82 Veit A, Wendisch VF & Polen T (2007) Global gene expression analysis of glucose overflow metabolism in *Escherichia coli* and reduction of aerobic acetate formation. *Appl. Microbiol. Biotechnol.* **74**, 406–421.

83 Waegeman H, Beauprez J, Moens H, Maertens J, De Mey M, Foulquié-Moreno MR, Heijnen JJ, Charlier D & Soetaert W (2011) Effect of *iclR* and *arcA* knockouts on biomass formation and metabolic fluxes in *Escherichia coli* K12 and its implications on understanding the metabolism of *Escherichia coli* BL21 (DE3). *BMC Microbiol.* **11**, 70.

84 El-Mansi M (2004) Flux to acetate and lactate excretions in industrial fermentations: Physiological and biochemical implications. *J. Ind. Microbiol. Biotechnol.* **31**, 295–300.

85 Zhuang K, Vemuri GN & Mahadevan R (2011) Economics of membrane occupancy and respiro-fermentation. *Mol. Syst. Biol.* **7**, 500.

86 De Mey M, Soetaert WK, De Maeseneire S & Vandamme E (2007) Minimizing acetate formation in *E. coli* fermentations. *J. Ind. Microbiol. Biotechnol.* **34**, 689 –700.

87 Castaño-Cerezo S, Pastor JM, Renilla S, Bernal V, Iborra JL & Cánovas M (2009) An insight into the role of phosphotransacetylase (*pta*) and the acetate/acetyl-CoA node in *Escherichia coli*. *Microb. Cell Fact.* **8**, 54.

88 Dittrich CR, Bennett GN & San K (2005) Characterization of the Acetate-Producing Pathways in *Escherichia coli*. *Biotechnol. Prog.* **21**, 1062–1067.

89 El-Mansi M & Holms WH (1989) Control of carbon flux to acetate excretion during growth of *Escherichia coli* in batch and continuous cultures. *J. Gen. Microbiol.* **135**, 2875–2883.

90 Yang YT, Bennett GN & San KY (1999) Effect of inactivation of *nuo* and *ackA-pta* on redistribution of metabolic fluxes in *Escherichia coli*. *Biotechnol. Bioeng.* **65**, 291–297.

91 Postma E, Verduyn C, Scheffers WA & Van Dijken JP (1989) Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **55**, 468–477.

92 Lin H, Bennett GN, Castro NM & San K-Y (2006) Acetyl-CoA synthetase overexpression in *Escherichia coli* demonstrates more efficient acetate assimilation and lower acetate accumulation: a potential tool in metabolic engineering. *Appl. Microbiol. Biotechnol.* **71**, 870–874.

93 Kumari S, Beatty CM, Browning DF, Busby SJW, Simel EJ, Wolfe AJ & Hovel-Miner G (2000) Regulation of acetyl coenzyme A synthetase in *Escherichia coli*. *J. Bacteriol.* **182**, 4173–4179.

94 Nanchen A, Schicker A, Sauer U & Revelles O (2008) Cyclic AMP-dependent catabolite repression is the dominant control mechanism of metabolic fluxes under glucose limitation in *Escherichia coli*. *J. Bacteriol.* **190**, 2323–2330.

95 Khankal R, Chin JW, Ghosh D & Cirino PC (2009) Transcriptional effects of CRP* expression in *Escherichia coli*. *J. Biol. Eng.* **3**, 13.

96 Sarkar D, Siddiquee KAZ, Araúzo-Bravo MJ, Oba T & Shimizu K (2008) Effect of *cra* gene knockout together with *edd* and *iclR* genes knockout on the metabolism in *Escherichia coli*. *Arch. Microbiol.* **190**, 559–571.

97 Scott M, Gunderson CW, Mateescu EM, Zhang Z & Hwa T (2010) Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science.* **330**, 1099–1102.

98 Flamholz A, Noor E, Bar-Even A, Liebermeister W & Milo R (2013) Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 10039–10044.

99 Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W & Selbach M (2011) Global quantification of mammalian gene expression control. *Nature* **473**, 337–342.

100 Marguerat S, Schmidt A, Codlin S, Chen W, Aebersold R & Bähler J (2012) Quantitative Analysis of Fission Yeast Transcriptomes and Proteomes in Proliferating and Quiescent Cells. *Cell* **151**, 671–683.

101 Gingold H & Pilpel Y (2011) Determinants of translation efficiency and accuracy. *Mol. Syst. Biol.* **7**, 481.

102 Lu P, Vogel C, Wang R, Yao X & Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* **25**, 117–124.

103 Lithwick G & Margalit H (2003) Hierarchy of sequence-dependent features associated with prokaryotic translation. *Genome Res.* **13**, 2665–2673.

104 Masuda T, Saito N, Tomita M & Ishihama Y (2009) Unbiased Quantitation of *Escherichia coli* Membrane Proteome Using Phase Transfer Surfactants. *Mol. Cell. Proteomics* **8**, 2770–2777.

105 Ishihama Y, Schmidt T, Rappsilber J, Mann M, Hartl FU, Kerner MJ & Frishman D (2008) Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics* **9**, 102.

106 Tuller T, Waldman YY, Kupiec M & Ruppin E (2010) Translation efficiency is determined by both codon bias and folding energy. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3645–3650.

107 Futcher B, Latter GI, Monardo P, Mclaughlin CS & Garrels I (1999) A Sampling of the Yeast Proteome. *Mol. Cell. Biol.* **19**, 7357–7368.

108 Maier T, Schmidt A, Güell M, Kühner S, Gavin A-C, Aebersold R & Serrano L (2011) Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol. Syst. Biol.* **7**, 511.

109 Trauchessec M, Jaquinod M, Bonvalot A, Brun V, Bruley C, Ropers D, De Jong H, Garin J, Bestel-Corre G & Ferro M (2014) Mass spectrometry-based workflow for accurate quantification of *E. coli* enzymes: how proteomics can play a key role in metabolic engineering. *Mol. Cell. Proteomics*, 954–968.

110 Taniguchi Y, Choi PJ, Li G-W, Chen H, Babu M, Hearn J, Emili A & Xie XS (2010) Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science.* **329**, 533–538.

111 Buescher JM, Liebermeister W, Jules M, Uhr M, Muntel J, Botella E, Hessling B, Kleijn RJ, Le Chat L, Lecointe F, Mader U, Nicolas P, Piersma S, Rugheimer F, Becher D, Bessieres P, Bidnenko E, Denham EL, Dervyn E, Devine KM, Doherty G, Drulhe S, Felicori L, Fogg MJ, Goelzer A, Hansen A, Harwood CR, Hecker M, Hubner S, Hultschig C, Jarmer H, Klipp E, Leduc A, Lewis P, Molina F, Noirot P, Peres S, Pigeonneau N, Pohl S, Rasmussen S, Rinn B, Schaffer M, Schnidder J, Schwikowski B, Van Dijl JM, Veiga P, Walsh S, Wilkinson AJ, Stelling J, Aymerich S & Sauer U (2012) Global Network Reorganization During Dynamic Adaptations of *Bacillus subtilis* Metabolism. *Science.* **335**, 1099–1103.

112 O'Brien EJ, Lerman JA, Chang RL, Hyduke DR & Palsson BØ (2013) Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.* **9**, 693.

113 Lerman JA, Hyduke DR, Latif H, Portnoy VA, Lewis NE, Orth JD, Schrimpe-Rutledge AC, Smith RD, Adkins JN, Zengler K & Palsson BØ (2012) *In silico* method for modelling metabolism and gene product expression at genome scale. *Nat. Commun.* **3**, 929.

114 Klumpp S, Zhang Z & Hwa T (2009) Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell* **139**, 1366–1375.

115 Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI & Covert MW (2012) A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell* **150**, 389–401.

116 Heinemann M & Sauer U (2010) Systems biology of microbial metabolism. *Curr. Opin. Microbiol.* **13**, 337–343.

117 Sauer U, Heinemann M & Zamboni N (2007) Genetics. Getting closer to the whole picture. *Science.* **316**, 550–551.

118 Yus E, Maier T, Michalodimitrakis K, van Noort V, Yamada T, Chen W-H, Wodke JAH, Güell M, Martínez S, Bourgeois R, Kühner S, Raineri E, Letunic I, Kalinina OV, Rode M, Herrmann R, Gutiérrez-Gallego R, Russell RB, Gavin A-C, Bork P & Serrano L (2009) Impact of genome reduction on bacterial metabolism and its regulation. *Science.* **326**, 1263–1268.

119 Güell M, van Noort V, Yus E, Chen W-H, Leigh-Bell J, Michalodimitrakis K, Yamada T, Arumugam M, Doerks T, Kühner S, Rode M, Suyama M, Schmidt S, Gavin A-C, Bork P & Serrano L (2009) Transcriptome complexity in a genome-reduced bacterium. *Science.* **326**, 1268–1271.

120 Kühner S, van Noort V, Betts MJ, Leo-Macias A, Batisse C, Rode M, Yamada T, Maier T, Bader S, Beltran-Alvarez P, Castaño-Diez D, Chen W-H, Devos D, Güell M, Norambuena T, Racke I, Rybin V, Schmidt A, Yus E, Aebersold R, Herrmann R, Böttcher B, Frangakis AS, Russell RB, Serrano L, Bork P & Gavin A-C (2009) Proteome organization in a genome-reduced bacterium. *Science.* **326**, 1235–1240.

121 Canelas AB, Harrison N, Fazio A, Zhang J, Pitkänen J-P, van den Brink J, Bakker BM, Bogner L, Bouwman J, Castrillo JI, Cankorur A, Chumnanpuen P, Daran-Lapujade P, Dikicioglu D, van Eunen K, Ewald JC, Heijnen JJ, Kirdar B, Mattila I, Mensonides FIC, Niebel A, Penttilä M, Pronk JT, Reuss M, Salusjärvi L, Sauer U, Sherman D, Siemann-Herzberg M, Westerhoff HV, de Winde J, Petranovic D, Oliver SG, Workman CT, Zamboni N & Nielsen J (2010) Integrated multilaboratory systems biology reveals differences in protein metabolism between two reference yeast strains. *Nat. Commun.* **1**, 145.

122 Gerosa L & Sauer U (2011) Regulation and control of metabolic fluxes in microbes. *Curr. Opin. Biotechnol.* **22**, 566–575.

123 Adamberg K, Seiman A & Vilu R (2012) Increased Biomass Yield of *Lactococcus lactis* by Reduced Overconsumption of Amino Acids and Increased Catalytic Activities of Enzymes. *PLoS One* **7**, e48223.

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

125 Van Eunen K, Rossell S, Bouwman J, Westerhoff HV & Bakker BM (2011) Quantitative analysis of flux regulation through hierarchical regulation analysis. *Methods Enzymol.* **500**, 571–595.

126 Kochanowski K, Sauer U & Chubukov V (2013) Somewhat in control - the role of transcription in regulating microbial metabolic fluxes. *Curr. Opin. Biotechnol.* **24**, 987–993.

127 Plotkin JB (2010) Transcriptional regulation is only half the story. *Mol. Syst. Biol.* **6**, 406.

128 Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV & Snoep JL (2000) Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* **267**, 5313–5329.

129 García-Contreras R, Vos P, Westerhoff HV & Boogerd FC (2012) Why *in vivo* may not equal *in vitro* - new effectors revealed by measurement of enzymatic activities under the same *in vivo*-like assay conditions. *FEBS J.* **279**, 4145–4159.

130 Maier T, Marcos J, Wodke JAH, Paetzold B, Liebeke M, Gutiérrez-Gallego R & Serrano L (2013) Large-scale metabolome analysis and quantitative integration with genomics and proteomics data in *Mycoplasma pneumoniae*. *Mol. Biosyst.* **9**, 1743–1755.

131 Sanghvi JC, Regot S, Carrasco S, Karr JR, Gutschow MV, Bolival B & Covert MW (2013) Accelerated discovery via a whole-cell model. *Nat. Methods* **10**, 1192–1195.

132 Bordbar A, Monk JM, King ZA & Palsson BØ (2014) Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* **15**, 107–120.

133 Macklin DN, Ruggero NA & Covert MW (2014) The future of whole-cell modeling. *Curr. Opin. Biotechnol.* **28**, 111–115.

134 Link H, Christodoulou D & Sauer U (2014) Advancing metabolic models with kinetic information. *Curr. Opin. Biotechnol.* **29**, 8–14.

135 Van Dien SJ (2013) From the first drop to the first truckload: commercialization of microbial processes for renewable chemicals. *Curr. Opin. Biotechnol.* **24**, 1061–1068.

136 Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL & Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008.

137 Kasemets K, Drews M, Nisamedtinov I, Paalme T & Adamberg K (2003) Modification of A-stat for the characterization of microorganisms. *J. Microbiol. Methods* **55**, 187–200.

138 Shinhara A, Matsui M, Hiraoka K, Nomura W, Hirano R & Nakahigashi K (2011) Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli*. *BMC Genomics* **12**, 428.

139 Yi H, Cho Y-J, Won S, Lee J-E, Jin Yu H, Kim S, Schroth GP, Luo S & Chun J (2011) Duplex-specific nuclease efficiently removes rRNA for prokaryotic RNA-seq. *Nucleic Acids Res.* **39**, e140.

140 Robertson B, Button D & Koch A (1998) Determination of the biomasses of small bacteria at low concentrations in a mixture of species with forward light scatter measurements by flow cytometry. *Appl. Environ. Microbiol.* **64**, 3900–3909.

141 Bratbak G & Dundas I (1984) Bacterial dry matter content and biomass estimations. *Appl. Environ. Microbiol.* **48**, 755–757.

142 Cayley S, Lewis BA, Guttman HJ & Record MT Jr (1991) Characterization of the cytoplasm of *Escherichia coli* K-12 as a function of external osmolarity. Implications for protein-DNA interactions *in vivo*. *J. Mol. Biol.* **222**, 281–300.

143 Arike L, Valgepea K, Peil L, Nahku R, Adamberg K & Vilu R (2012) Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *J. Proteomics* **75**, 5437–5448.

144 Lowry O, Rosebrough N, Lewis Farr A & Randall R (1951) Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265–275.

145 Adamberg K, Lahtvee P-J, Valgepea K, Abner K & Vilu R (2009) Quasi steady state growth of *Lactococcus lactis* in glucose-limited acceleration stat (A-stat) cultures. *Antonie Van Leeuwenhoek* **95**, 219–226.

146 Benjamini Y & Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Statist. Soc.* **B57**, 289–300.

147 Jensen KF (1993) The *Escherichia coli* K-12 “wild types” W3110 and MG1655 have an *rph* frameshift mutation that leads to pyrimidine starvation due to low *pyrE* expression levels. *J. Bacteriol.* **175**, 3401–3407.

148 Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V & Palsson BØ (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* **3**, 121.

149 Lee MV, Topper SE, Hubler SL, Hose J, Wenger CD, Coon JJ & Gasch AP (2011) A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. *Mol. Syst. Biol.* **7**, 514.

150 Lackner DH, Schmidt MW, Wu S, Wolf DA & Bahler J (2012) Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. *Genome Biol.* **13**, R25.

151 Deutscher J, Postma PW & Francke C (2006) How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* **70**, 939–1031.

152 Escalante A, Salinas Cervantes A, Gosset G & Bolívar F (2012) Current knowledge of the *Escherichia coli* phosphoenolpyruvate-carbohydrate phosphotransferase system: peculiarities of regulation and impact on growth and product formation. *Appl. Microbiol. Biotechnol.* **94**, 1483–1394.

153 Barak R, Abouhamad WN & Eisenbach M (1998) Both acetate kinase and acetyl coenzyme A synthetase are involved in acetate-stimulated change in the direction of flagellar rotation in *Escherichia coli*. *J. Bacteriol.* **180**, 985–988.

154 Klein A, Shulla A, Reimann SA, Keating D & Wolfe AJ (2007) The intracellular concentration of acetyl phosphate in *Escherichia coli* is sufficient for direct phosphorylation of two-component response regulators. *J. Bacteriol.* **189**, 5574–5581.

155 Mayover TL, Halkides CJ & Stewart RC (1999) Kinetic characterization of CheY phosphorylation reactions: comparison of P-CheA and small-molecule phosphodonors. *Biochemistry* **38**, 2259–2271.

156 Weinert BT, Iesmantavicius V, Wagner SA, Schölz C, Gummesson B, Beli P, Nyström T & Choudhary C (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. *Mol. Cell* **51**, 265–272.

157 Bernal V, Castaño-Cerezo S, Gallego-Jara J, Ecija-Conesa A, de Diego T, Iborra JL & Cánovas M (2014) Regulation of bacterial physiology by lysine acetylation of proteins. *N. Biotechnol.* DOI: 10.1016/j.nbt.2014.03.002.

158 Kuhn ML, Zemaitaitis B, Hu LI, Sahu A, Sorensen D, Minasov G, Lima BP, Scholle M, Mrksich M, Anderson WF, Gibson BW, Schilling B & Wolfe AJ (2014) Structural, kinetic and proteomic characterization of acetyl phosphate-dependent bacterial protein acetylation. *PLoS One* **9**, e94816.

159 Anfora AT, Halladin DK, Haugen BJ & Welch RA (2008) Uropathogenic *Escherichia coli* CFT073 is adapted to acetatogenic growth but does not require acetate during murine urinary tract infection. *Infect. Immun.* **76**, 5760–5767.

160 Mizrahi I, Biran D & Ron EZ (2009) Involvement of the Pta-AckA pathway in protein folding and aggregation. *Res. Microbiol.* **160**, 80–84.

161 Wolfe AJ, Chang D-E, Walker JD, Seitz-Partridge JE, Vidaurri MD, Lange CF, Prüss BM, Henk MC, Larkin JC & Conway T (2003) Evidence that acetyl phosphate functions as a global signal during biofilm development. *Mol. Microbiol.* **48**, 977–988.

162 Treves DS, Manning S & Adams J (1998) Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*. *Mol. Biol. Evol.* **15**, 789–797.

163 Castaño-Cerezo S, Bernal V, Blanco-Catalá J, Iborra JL & Cánovas M (2011) cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in *Escherichia coli*. *Mol. Microbiol.* **82**, 1110–1128.

164 Zhao K, Chai X & Marmorstein R (2004) Structure and substrate binding properties of cobB, a Sir2 homolog protein deacetylase from *Escherichia coli*. *J. Mol. Biol.* **337**, 731–741.

165 Covert MW, Knight EM, Reed JL, Herrgard MJ & Palsson BØ (2004) Integrating high-throughput and computational data elucidates bacterial networks. *Nature* **429**, 92–96.

166 Perrenoud A & Sauer U (2005) Impact of global transcriptional regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on glucose catabolism in *Escherichia coli*. *J. Bacteriol.* **187**, 3171–3179.

167 Otto A, Bernhardt J, Meyer H, Schaffer M, Herbst F-A, Siebourg J, Mäder U, Lalk M, Hecker M & Becher D (2010) Systems-wide temporal proteomic profiling in glucose-starved *Bacillus subtilis*. *Nat. Commun.* **1**, 137.

168 Schaechter M, Maaløe O & Kjeldgaard NO (1958) Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of *Salmonella typhimurium*. *J. Gen. Microbiol.* **19**, 592–606.

169 Raser JM & O'Shea EK (2005) Noise in gene expression: origins, consequences, and control. *Science* **309**, 2010–2013.

170 Sharp PM & Li W (1987) The codon adaptation index - a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **15**, 1281–1295.

171 Kudla G, Murray A, Tollervey D & Plotkin JB (2009) Coding-Sequence Determinants of Gene Expression in *Escherichia coli*. *Science.* **324**, 255–258.

172 Supek F & Šmuc T (2010) On relevance of codon usage to expression of synthetic and natural genes in *Escherichia coli*. *Genetics* **185**, 1129–1134.

173 Vogel C, Abreu RDS, Ko D, Le S-Y, Shapiro BA, Burns SC, Sandhu D, Boutz DR, Marcotte EM & Penalva LO (2010) Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. *Mol. Syst. Biol.* **6**, 400.

174 Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ & Natale DA (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**, 41.

175 Tobias JW, Shrader TE, Rocap G & Varshavsky A (1991) The N-end rule in bacteria. *Science.* **254**, 1374–1377.

176 Bernstein JA, Khodursky AB, Lin P-H, Lin-Chao S & Cohen SN (2002) Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 9697–9702.

177 Selinger DW, Saxena RM, Cheung KJ, Church GM & Rosenow C (2003) Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.* **13**, 216–223.

178 Zhou Y, Vazquez A, Wise A, Warita T, Warita K, Bar-Joseph Z & Oltvai ZN (2013) Carbon catabolite repression correlates with the maintenance of near invariant molecular crowding in proliferating *E. coli* cells. *BMC Syst. Biol.* **7**, 138.

179 Ghaemmaghami S, Huh W-K, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK & Weissman JS (2003) Global analysis of protein expression in yeast. *Nature* **425**, 737–741.

180 Schmidt A, Beck M, Malmström J, Lam H, Claassen M, Campbell D & Aebersold R (2011) Absolute quantification of microbial proteomes at different states by directed mass spectrometry. *Mol. Syst. Biol.* **7**, 510.

181 Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, Szymborska A, Herzog F, Rinner O, Ellenberg J & Aebersold R (2011) The quantitative proteome of a human cell line. *Mol. Syst. Biol.* **7**, 549.

182 Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, Pääbo S & Mann M (2011) Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* **7**, 548.

183 Akashi H & Gojobori T (2002) Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 3695–3700.

184 Jin DJ, Cagliero C & Zhou YN (2011) Growth rate regulation in *Escherichia coli*. *FEMS Microbiol. Rev.* **36**, 269–287.

185 Oleinikov AV, Jokhadze GG & Traut RR (1998) A single-headed dimer of *Escherichia coli* ribosomal protein L7/L12 supports protein synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 4215–4218.

186 Haverkorn van Rijsewijk BRB, Nanchen A, Nallet S, Kleijn RJ & Sauer U (2011) Large-scale ^{13}C-flux analysis reveals distinct transcriptional control of respiratory and fermentative metabolism in *Escherichia coli*. *Mol. Syst. Biol.* **7**, 477.

187 Wessely F, Bartl M, Guthke R, Li P, Schuster S & Kaleta C (2011) Optimal regulatory strategies for metabolic pathways in *Escherichia coli* depending on protein costs. *Mol. Syst. Biol.* **7**, 515.

188 Lourenço A, Carneiro S, Pinto JP, Rocha M, Ferreira EC & Rocha I (2011) A study of the short and long-term regulation of *E. coli* metabolic pathways. *J. Integr. Bioinform.* **8**, 183.

189 Chubukov V, Gerosa L, Kochanowski K & Sauer U (2014) Coordination of microbial metabolism. *Nat. Rev. Microbiol.* **12**, 327–340.

190 Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS & Milo R (2011) The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410.

191 Nam H, Lewis NE, Lerman JA, Lee D-H, Chang RL, Kim D & Palsson BØ (2012) Network Context and Selection in the Evolution to Enzyme Specificity. *Science.* **337**, 1101–1104.

192 Taymaz-Nikerel H, van Gulik WM & Heijnen JJ (2011) *Escherichia coli* responds with a rapid and large change in growth rate upon a shift from glucose-limited to glucose-excess conditions. *Metab. Eng.* **13**, 307–318.

193 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 (2007) The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 15753–15758.

194 Young R & Bremer H (1976) Polypeptide-chain-elongation rate in *Escherichia coli* B/r as a function of growth rate. *Biochem. J.* **160**, 185–194.

PUBLICATION I

Valgepea K, Adamberg K, Nahku R, Lahtvee P-J, Arike L & Vilu R

Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase.

BMC Syst. Biol. 4, 166, (2010)

Valgepea et al. BMC Systems Biology 2010, 4:166
http://www.biomedcentral.com/1752-0509/4/166

BMC Systems Biology

RESEARCH ARTICLE Open Access

Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase

Kaspar Valgepea[1,2], Kaarel Adamberg[2,3], Ranno Nahku[1,2], Petri-Jaan Lahtvee[1,2], Liisa Arike[2,3], Raivo Vilu[1,2*]

Abstract

Background: The biotechnology industry has extensively exploited *Escherichia coli* for producing recombinant proteins, biofuels etc. However, high growth rate aerobic *E. coli* cultivations are accompanied by acetate excretion *i.e.* overflow metabolism which is harmful as it inhibits growth, diverts valuable carbon from biomass formation and is detrimental for target product synthesis. Although overflow metabolism has been studied for decades, its regulation mechanisms still remain unclear.

Results: In the current work, growth rate dependent acetate overflow metabolism of *E. coli* was continuously monitored using advanced continuous cultivation methods (A-stat and D-stat). The first step in acetate overflow switch (at $\mu = 0.27 \pm 0.02$ h^{-1}) is the repression of acetyl-CoA synthethase (Acs) activity triggered by carbon catabolite repression resulting in decreased assimilation of acetate produced by phosphotransacetylase (Pta), and disruption of the PTA-ACS node. This was indicated by acetate synthesis pathways PTA-ACKA and POXB component expression down-regulation before the overflow switch at $\mu = 0.27 \pm 0.02$ h^{-1} with concurrent 5-fold stronger repression of acetate-consuming Acs. This in turn suggests insufficient Acs activity for consuming all the acetate produced by Pta, leading to disruption of the acetate cycling process in PTA-ACS node where constant acetyl phosphate or acetate regeneration is essential for *E. coli* chemotaxis, proteolysis, pathogenesis etc. regulation. In addition, two-substrate A-stat and D-stat experiments showed that acetate consumption capability of *E. coli* decreased drastically, just as Acs expression, before the start of overflow metabolism. The second step in overflow switch is the sharp decline in cAMP production at $\mu = 0.45$ h^{-1} leading to total Acs inhibition and fast accumulation of acetate.

Conclusion: This study is an example of how a systems biology approach allowed to propose a new regulation mechanism for overflow metabolism in *E. coli* shown by proteomic, transcriptomic and metabolomic levels coupled to two-phase acetate accumulation: acetate overflow metabolism in *E. coli* is triggered by Acs down-regulation resulting in decreased assimilation of acetic acid produced by Pta, and disruption of the PTA-ACS node.

Background

Escherichia coli has not only been the prime organism for developing new molecular biology methods but also for producing recombinant proteins, low molecular weight compounds etc. in industrial biotechnology for decades due to its low cost manufacturing and end-product purification and its ability to reach high cell densities grown aerobically [1,2]. However, a major problem exists with aerobic *E. coli* cultivation on glucose at high growth rates-formation and accumulation of considerable amounts of acetic acid *i.e.* overflow metabolism. In addition to being detrimental for target product synthesis, accumulated acetate inhibits growth and diverts valuable carbon from biomass formation [3,4].

The acetate synthesis and utilization pathways [5] can be seen in Figure 1: acetate can be synthesized by

* Correspondence: raivo@kbfi.ee
[1]Tallinn University of Technology, Department of Chemistry, Akadeemia tee 15, 12618 Tallinn, Estonia
Full list of author information is available at the end of the article

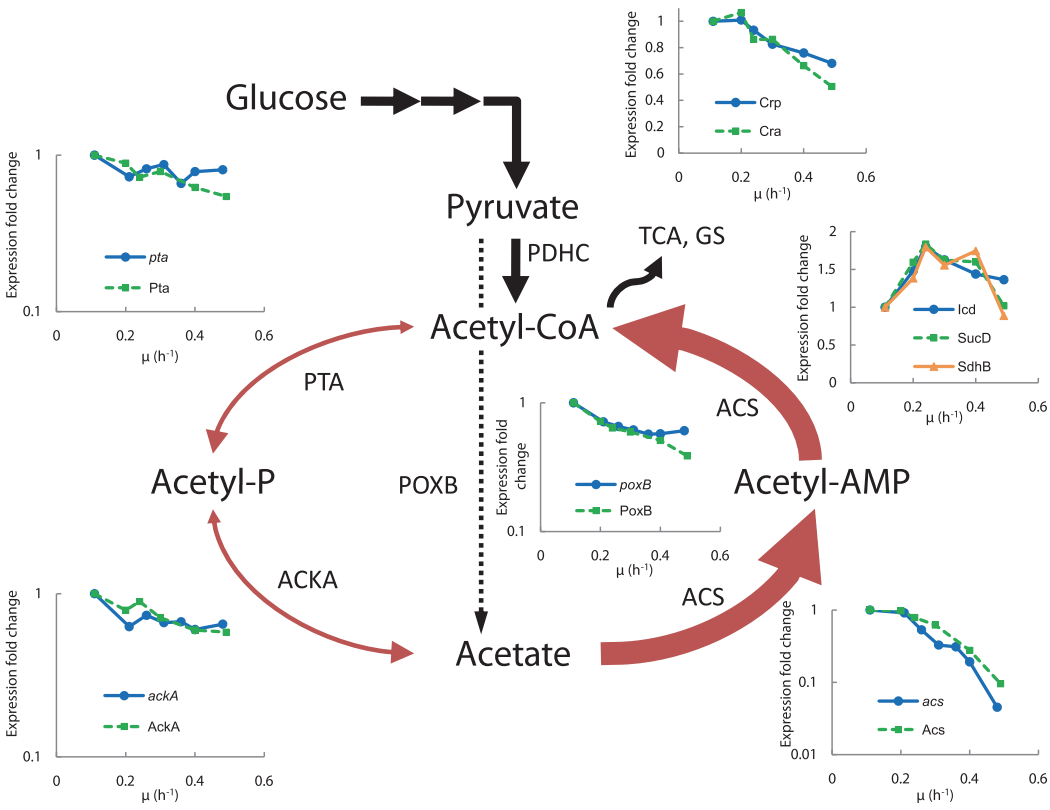

Figure 1 Effect of specific growth rate on acetate synthesis and utilization pathways, selected TCA cycle and carbon catabolite repressed gene and protein expression levels in *E. coli* A-stat experiments. PTA, phosphotransacetylase; ACKA, acetate kinase; ACS, acetyl-CoA synthetase; POXB, pyruvate oxidase; PDHC, pyruvate dehydrogenase complex; TCA, tricarboxylic acid cycle; GS, glyoxylate shunt; Crp, cyclic AMP receptor protein; Cra, catabolite repressor activator; Icd, isocitrate dehydrogenase; SucD, succinyl-CoA synthethase; SdhB, succinate dehydrogenase; μ, specific growth rate (h^{-1}). Thickness of red arrows denotes level of ACS and PTA-ACKA pathway repression (thick line represents stronger repression). Protein data points are average of two independent experiments, error bars are not shown for better visualization. Gene names are italicized. Refer to Additional file 2 for standard deviations and all the data.

phosphotransacetylase (PTA)/acetate kinase (ACKA) and by pyruvate oxidase (POXB). Acetic acid can be metabolized to acetyl-CoA either by the PTA-ACKA pathway or by acetyl-CoA synthetase (ACS) through an intermediate acetyl-AMP. The high affinity (K_m of 200 μM for acetic acid) ACS scavenges acetate at low concentrations whereas the low affinity PTA-ACKA pathway (K_m of 7-10 mM) is activated in the presence of high acetate concentrations [6].

The phenomenon of overflow metabolism has been studied widely over the years and it is commonly believed to be caused by an imbalance between the fluxes of glucose uptake and those for energy production and biosynthesis [7,8]. Several explanations such as the saturation of catalytic activities in the tricarboxylic acid (TCA) cycle [9,10] and respiratory chain [7,11,12], energy generation [5,13] or the necessity for coenzyme A replenishment [14] have been proposed. In addition to bioprocess level approaches [1,15], various genetic modifications of the acetate synthesis pathways extensively reviewed in De Mey *et al.* [15] have been made to minimize acetic acid production. For instance, it has been shown that deleting the main acetate synthesis route PTA-ACKA results in a strong reduction (up to 80%) of acetate excretion, maximum growth rate (*ca* 20%) and elevated levels of formate and lactate (*ca* 30-fold) [4,16-18], whereas *poxB* disruption causes reduction in biomass yield (*ca* 25%) and loss of aerobic growth efficiency of *E. coli* [19]. The latter indicates that acetate excretion cannot be simply excluded by

disrupting its synthesis routes without encountering other unwanted effects. Unfortunately, no clear conclusions could be drawn from batch experiments with an *acs* knock-out strain [4]. It should be noted that studies with *E. coli* genetically modified strains engineered to diminish acetate production in batch cultures have not fully succeeded in avoiding acetate accumulation together with increasing target product production yields and rates [15]. Additionally, these studies have not allowed elucidating the mechanism of overflow metabolism unequivocally [4,20,21].

Acetate overflow is a growth rate dependent phenomenon, but no study has specifically focused on growth rate dependency of protein and gene expression regulation, intra-and extracellular metabolite levels using also metabolic modeling. Describing the physiology of an organism on several 'omic levels is the basis of systems biology that facilitates better understanding of metabolic regulation [22]. In this study, *E. coli* metabolism at proteomic, transcriptomic and metabolomic levels was investigated using continuous cultivation methods prior to and after overflow metabolism was switched on. Usually, chemostat cultures are used for steady state metabolism analysis, however, we applied two changestat cultivation techniques: accelerostat (A-stat) and dilution rate stat (D-stat), see Methods section for details [23,24]. These cultivation methods were used as they provide three advantages over chemostat. Firstly, these changestat cultivation techniques precisely detect metabolically relevant switch points (*e.g.* start of overflow metabolism, maximum specific growth rate) and enable to monitor the dynamic patterns of several metabolic physiological responses simultaneously which could be left unnoticed using chemostat. Secondly, it is possible to collect vast amount of steady state comparable samples and by doing so, save time. Thirdly, both A-stat and D-stat enable to quantitatively study specific growth rate dependent co-utilization of growth substrates. Latter advantage was applied for investigating acetic acid consumption capability of *E. coli* at various dilution rates in this study. Combining changestat cultivation methods enables to study metabolism responses of the same genotype at different physiological states in detail without encountering the possible metabolic artifacts accompanied when using genetically modified strains.

Results obtained by studying specific growth rate dependent changes in *E. coli* proteome, transcriptome and metabolome in continuous cultures together with metabolic modeling allowed us to propose a new theory for acetate overflow: acetate excretion in *E. coli* is triggered by carbon catabolite repression mediated down-regulation of Acs resulting in decreased assimilation of acetate produced by Pta, and disruption of the PTA-ACS node.

Results

E. coli metabolic switch points characterization

In all accelerostat (A-stat) cultivation experiments, after the culture had been stabilized in chemostat at 0.10 h^{-1} to achieve steady state conditions, continuous increase in dilution rate with acceleration rate (a) 0.01 h^{-2} (0.01 h^{-1} per hour) was started. Continuous change of specific growth rate resulted in detecting several important changes in *E. coli* metabolism as demonstrated in Figure 2. Firstly, in A-stat cultivations where glucose was the only carbon source in the medium, acetic acid started to accumulate (*i.e.* overflow metabolism switch) at μ = 0.27 ± 0.02 h^{-1} (average ± standard deviation) and a two-phase acetate accumulation pattern was observed (discussed below; Figure 2). Cells reached maximum CO_2 production and O_2 consumption at μ = 0.46 ± 0.02 h^{-1} and metabolic fluctuations were observed at μ = 0.49 ± 0.03 h^{-1} followed by washout of culture at μ = 0.54 ± 0.03 h^{-1} (corresponding to maximum specific growth rate at given conditions). The nature of these fluctuations will be studied further and not covered in the current publication. All A-stat results were reproduced with relative standard deviation less than 10% with the exception of acetate production per biomass (Y_{OAc^-}) (Table 1 and Figure S1 in Additional file 1).

Metabolomic responses to rising specific growth rate

A-stat cultivation enabled to study acetic acid accumulation profile in detail with increasing specific growth rate. Interestingly, a two-phase acetate accumulation pattern was observed (Figure 2). Slow accumulation of acetic

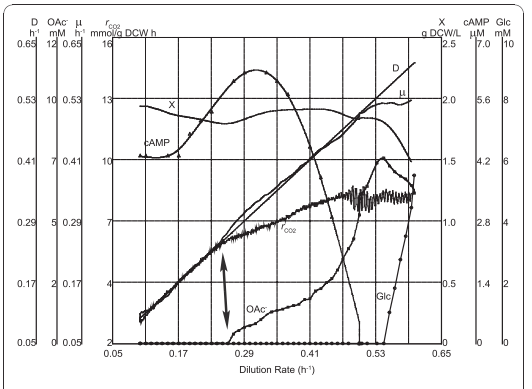

Figure 2 Increasing dilution rate dependent *E. coli* metabolism characterization in one A-stat cultivation (a = 0.01 h^{-2}). D, dilution rate (h^{-1}); X, biomass concentration (g dry cellular weight (DCW)/L); μ, specific growth rate (h^{-1}); r_{CO2}, specific CO_2 production rate (mmol/g DCW h); OAc^-, acetate concentration (mM); Glc, glucose concentration (mM); cAMP, cyclic AMP concentration (μM). Arrow indicates the start of overflow metabolism. Start of vertical axes was chosen for better visualization.

Table 1 A-stat and chemostat growth characteristics comparison and A-stat reproducibility over the studied specific growth rate range for three independent experiments

	μ = 0.24 h^{-1}		μ = 0.30 h^{-1}		μ = 0.40 h^{-1}		μ = 0.51 h^{-1}		μ = 0.10-0.47 h^{-1}
	Chemostat	A-stat	Chemostat	A-stat	Chemostat	A-stat	Chemostat	A-stat	A-stat RSD, %
Y_{XS}[a]	0.44	0.40 ± 0.01	0.46	0.41 ± 0.01	0.44	0.42 ± 0.00	0.43	0.41 ± 0.01	2.0
Y_{OAc^-}[b]	NDE	NDE	0.53	0.90 ± 0.32	1.70	1.56 ± 0.23	3.25	3.35 ± 0.82	ND
Y_{cAMP}[c]	3.47	3.59 ± 0.39	3.25	3.55 ± 0.32	2.70	2.17 ± 0.07	0.86	0.71[e]	9.1
Y_{CO2}[d]	27.56	30.12 ± 2.04	27.55	27.19 ± 1.22	26.24	23.86 ± 1.41	ND	21.19 ± 0.19	5.6

A-stat values represent the average from three independent experiments and standard deviation follows the ± sign. Chemostat values from one experiment. NDE, not detected. ND, not determined. RSD, relative standard deviation.

[a]Biomass yield is given in g dry cell weight (DCW)/g glucose consumed (g DCW/g glucose).

[b]Acetic acid production per biomass is given in mmol acetic acid/g DCW.

[c]cAMP production per biomass is given in μmol cAMP/g DCW.

[d]Carbon dioxide (CO_2) production per biomass is given in mmol CO_2/g DCW.

[e]Data from one A-stat experiment.

acid started at μ = 0.27 ± 0.02 h^{-1} with concomitant change in specific CO_2 production rate (Figure 2). Faster accumulation of acetate was witnessed after cells had reached maximum CO_2 production at μ = 0.46 ± 0.02 h^{-1}. Quite surprisingly, production of the important carbon catabolite repression (CCR) signal molecule cAMP (Y_{cAMP}) rose from steady state chemostat level 2.45 ± 0.26 μmol/g dry cellular weight (DCW) (μ = 0.10 h^{-1}) to 3.55 ± 0.32 μmol/g DCW (μ = 0.30 h^{-1}) after which it sharply decreased to 1.30 ± 0.44 μmol/g DCW at μ = 0.45 h^{-1} (Figure S1 in Additional file 1). This abrupt decline took place simultaneously with the faster acetate accumulation profile described above (Figure 2 and Figure S1 in Additional file 1). In addition, similar two-phase acetate accumulation phenomenon was observed in a two-substrate (glucose + acetic acid) A-stat during the decrease of cAMP around specific growth rate 0.39 h^{-1} (Figure S2 in Additional file 1).

Significant fall in two of the measured pentose phosphate pathway intermediates ribose-5-phosphate (R5P) and erythrose-4-phosphate (E4P) was detected with increasing specific growth rate which could point to possible limitation in RNA biosynthesis during growth (Figure 3A). PTA-ACS node related compound nonesterified acetyl-CoA (HS-CoA) level declined two-fold simultaneously with cAMP after acetate started to accumulate (Figure 3B). This indicates the possible increase of other CoA containing compounds *e.g.* succinyl-CoA. Accumulation of TCA cycle intermediates α-ketoglutarate and isocitrate (Figure 3B) with increasing dilution rate could be associated with pyrimidine deficiency and decrease of ATP expenditure in the PTA-ACS cycle. Concurrently, intracellular concentrations of fructose-1,6-bisphosphate (FBP) and glyceraldehyde-3-phosphate (GAP) from the upper part of energy generating glycolysis increased 6- and 3-fold, respectively (Figure 3C).

Functional-genomic responses to rising specific growth rate

The two main known pathways for acetate synthesis phosphotransacetylase-acetate kinase (PTA-ACKA) and pyruvate oxidase (POXB) were down-regulated, both on gene and protein expression levels, from μ = 0.20 h^{-1} *i.e.* before acetate overflow was switched on. At the same time, there was a concurrent 10-fold repression of the acetic acid utilization enzyme acetyl-CoA synthetase (Acs). This substantial difference (5-fold) between the acetate synthesis and assimilation pathways expression suggests that the synthesized acetic acid cannot be fully assimilated with increasing growth rates (Figure 1).

We observed the beginning of carbon catabolite repression (CCR) induction prior to acetate accumulation in parallel with Acs down-regulation. This was indicated by down-regulation (3-fold on average) of CCR-mediated components: alternative (to glucose) substrate transport and utilization systems like galactose (MglAB), maltose (MalBEFKM), galactitol (GatABC), L-arabinose (AraF), D-ribose (RbsAB), C_4-dicarboxylates (DctA) and acetate (ActP, YjcH) (Figure 4C and Additional file 2). Moreover, expression of transcription activator Crp (cyclic AMP receptor protein which regulates the expression of Acs transcribing *acs-yjcH-actP* operon) and Cra (catabolite repressor activator; a global transcriptional protein essential for acetic acid uptake [25]) were reduced 1.5 and 2 times, respectively, in like manner to carbon catabolite repressed proteins mentioned above (Figure 1). Simultaneously, components of the gluconeogenesis pathway (Pck, MaeB, Pps) and glyoxylate shunt enzymes AceA, AceB (vital for acetate consumption) were repressed with growth rate increase (Figure 4B and Additional file 2). It should be emphasized that most of the TCA cycle gene and protein levels were maintained or even increased up to μ = 0.40 h^{-1}

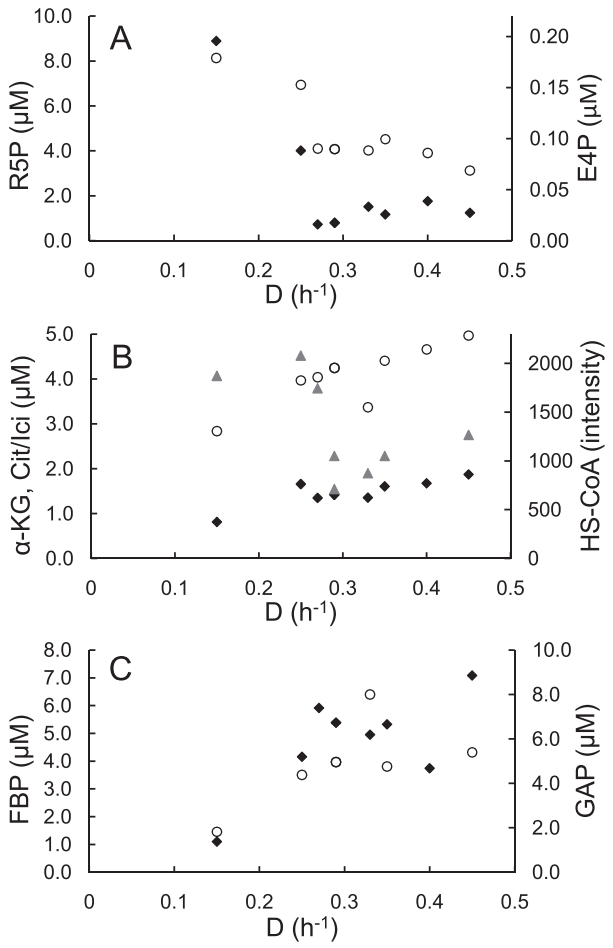

Figure 3 Dilution rate dependent intracellular metabolite patterns in one *E. coli* A-stat experiment. D, dilution rate (h^{-1}). (A) Pentose phosphate pathway metabolites. R5P, ribose-5-phosphate concentration (black diamond); E4P, erythrose-4-phosphate concentration (open circle). (B) TCA cycle metabolites and co-factor free CoA. α-KG, α-ketoglutarate concentration (black diamond); Cit/Ici, citrate/isocitrate pool concentration (open circle); HS-CoA, co-factor free CoA level (grey triangle). (C) Glycolysis (upper part) metabolites. FBP, fructose-1,6-bisphosphate concentration (black diamond); GAP, glyceraldehyde-3-phosphate concentration (open circle).

followed by sudden repression simultaneous to achieving maximum specific CO_2 production rate (μ = 0.46 ± 0.02 h^{-1}, see above; Figure 1 Figure 2 and Figure 4A). This may allude to no limitation at the TCA cycle level around the specific growth rate where overflow metabolism was switched on.

Acetic acid consumption capability studied by dilution rate stat (D-stat) and two-substrate A-stat cultivations

The beginning of a strong decrease in acetate assimilation enzyme Acs expression before overflow switch point implies to a possible connection between acetate assimilation capability and overflow metabolism of acetate (Figure 1). Therefore, specific growth rate dependent acetic acid consumption capabilities were investigated using D-stat and two-substrate A-stat methods. It was shown by D-stat experiments at various dilution rates that more than a 12-fold reduction in acetate consumption capability took place around overflow switch point, and its total loss was detected between dilution rates 0.45 and 0.505 ± 0.005 h^{-1} (Figure 5). Acetic acid consumption and production was also studied in a single experiment using two substrate (glucose + acetic acid) A-stat cultivation (Figure S2 in Additional file 1) which

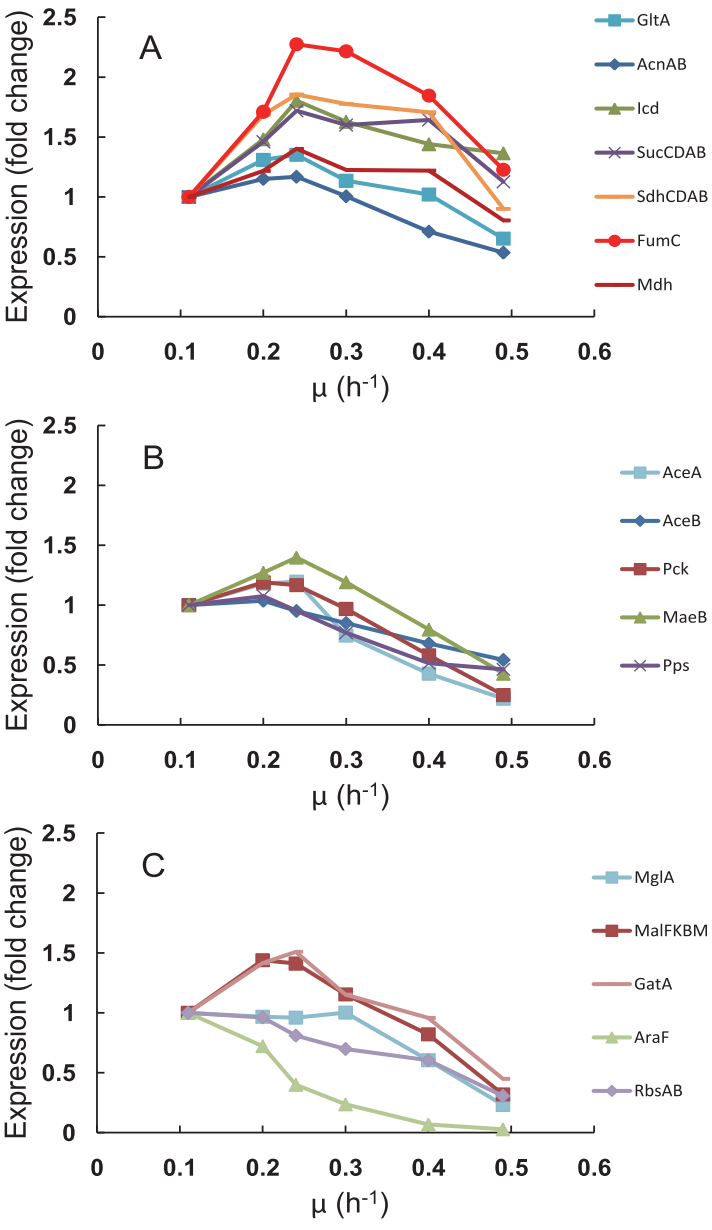

Figure 4 Specific growth rate dependent TCA cycle, glyoxylate shunt, glyconeogenesis and carbon catabolite repressed protein expression changes in *E coli* A-stat cultures. μ, specific growth rate (h^{-1}). (A) TCA cycle (average of proteins from the same operon are depicted as one point *e.g.* AcnAB). (B) Glyoxylate shunt (AceA, AceB) and glyconeogenesis. (C) Carbon catabolite repressed proteins. Protein data points are average of two independent experiments, error bars are not shown for better visualization (refer to Additional file 2 for standard deviations).

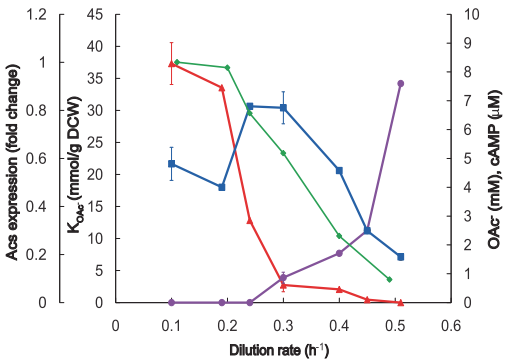

Figure 5 Dilution rate dependent acetic acid consumption capability in *E. coli* D-stat cultures. K_{OAc^-}, acetic acid consumption per biomass (red triangle); OAc^-, acetate concentration in chemostat before the start of acetic acid supplemented medium addition (violet circle); cAMP, cyclic AMP concentration (blue square); Acs, acetyl-CoA synthetase protein dilution rate dependent expression levels from A-stat (green diamond). Error bars represent the standard deviation from two independent D-stat experiments.

demonstrated that acetic acid consumption started to decrease at $\mu = 0.25\ h^{-1}$ and was completely abolished at $\mu = 0.48\ h^{-1}$ which fits into the range of dilution rates observed in D-stat.

A-stat comparison with chemostat

As could be seen from Table 1 major growth characteristics such as biomass yield (Y_{XS}), acetate (Y_{OAc^-}), cyclic AMP (Y_{cAMP}) and carbon dioxide (Y_{CO2}) production per biomass from A-stat and chemostat are all fully quantitatively comparable. The latter results enable to use A-stat data for quantitative modeling calculations. In addition, the two continuous cultivation methods were examined at transcriptome level using DNA microarrays. Transcript spot intensities from quasi steady state A-stat sample at $\mu = 0.48\ h^{-1}$ and chemostat sample at $\mu = 0.51\ h^{-1}$ showed an excellent Pearson product-moment correlation coefficient R = 0.964 (Figure S3 in Additional file 1; Additional file 3). This indicates good biological correlation between *E. coli* transcript profiles at similar specific growth rates in chemostat and A-stat. These results showed that our quasi steady state data from A-stat and D-stat cultures are steady state representative.

Proteome and transcriptome comparison

E. coli protein expression ratios for around 1600 proteins were generated by comparing two biological replicates at specific growth rates 0.20 ± 0.01; 0.26; 0.30 ± 0.01; 0.40 ± 0.00; 0.49 ± 0.01 h^{-1} with sample at $\mu = 0.10 \pm 0.01\ h^{-1}$ (chemostat point prior to the start of acceleration in A-stat) which produced Pearson correlation coefficients for two biological replicates in the indicated pairs of comparison in the range of R = 0.788-0.917 (Figure S4 in Additional file 1).

DNA microarray analysis of 4,321 transcripts was conducted with the Agilent platform using the samples from one A-stat cultivation. Gene expression ratios between specific growth rates 0.21; 0.26; 0.31; 0.36; 0.40; 0.48 h^{-1} and $\mu = 0.11\ h^{-1}$ (chemostat point prior to the start of acceleration in A-stat) were calculated. Comparison of gene and protein expression changes (between respective specific growth rates) revealed that components of the PTA-ACS node were regulated at transcriptional level as the absolute majority of the studied transcripts and proteins indicated by the good correlation between transcriptome and proteome expression profiles (Figure 1 and Figure S5 in Additional file 1).

Most recent studies have either failed to find a significant correlation between protein and mRNA abundances or have observed only a weak correlation (reviewed in [22]). It has been suggested that the main reasons for uncoupling of mRNA and protein abundances are protein regulation by post-translational modification, post-transcriptional regulation of protein synthesis, differences in the half-lives of mRNA and proteins, or possible functional requirement for protein binding [22]. As the cells in these studies were mostly cultured in non steady state condition, our steady state data with very good correlation between transcriptome and proteome implies that the physiological state of the culture (steady state vs. non steady state) could be an important factor in terms of mRNA and protein correlation determination. Transcriptome and proteome data are presented in Additional file 2 and at NCBI Gene Expression Omnibus and PRIDE database (see Methods for details), respectively.

Discussion

To gain better insights into the regulation of acetate overflow metabolism in *E. coli*, we studied specific growth rate dependent proteomic, transcriptomic and metabolomic patterns combined with metabolic modeling using advanced continuous cultivation methods, which has not been carried out before. Continuous monitoring of the specific growth rate effect on *E. coli* metabolism enabled us to precisely detect important metabolic shift points, the most important being the start of acetate overflow at $\mu = 0.27 \pm 0.02\ h^{-1}$ (Figure 2), and changing patterns of a number of important metabolites *e.g.* acetate, cAMP. Quite surprising was the down-regulation of the known acetate synthesis pathways, PTA-ACKA and POXB expression before overflow switch with increasing growth rate (Figure 1). A similar pattern has been seen before in chemostat cultures

but without emphasizing the possible physiological consequences [26-28]. A 10-fold repression of the acetic acid utilization enzyme acetyl-CoA synthetase (Acs) expression was observed concurrently with the down-regulation of the PTA-ACKA pathway indicating that acetic acid synthesis might exceed its assimilation (Figure 1). Our two substrate A-stat and D-stat experiments directly proved that acetate consumption capability of *E. coli* is specific growth rate dependent as acetate consumption started to decrease at $\mu = 0.25\ h^{-1}$ (Figure S2 in Additional file 1) and acetate consumption capability decreased 12-fold around overflow switch growth rate $\mu = 0.27 \pm 0.02\ h^{-1}$, respectively (Figure 5). In addition, it was shown that activation of carbon catabolite repression (CCR) and repression of Acs take place simultaneously prior to the start of overflow metabolism (Figure 1 Figure 4 and Figure 5). As a result, it is proposed that acetate overflow metabolism in *E. coli* is triggered by Acs down-regulation resulting in decreased assimilation of acetic acid produced by Pta, and disruption of the PTA-ACS node.

We showed that Acs was concurrently down-regulated five times more compared to the acetate synthesis pathways (Figure 1). In addition, the TCA cycle flux decrease as shown by change in CO_2 production at overflow switch growth rate indicates that carbon is not metabolized by the TCA cycle after the start of acetate accumulation with pre overflow switch rates (Figure 2 and Additional file 4). The latter is caused because the amount of acetyl-CoA entering the TCA cycle decreases after carbon is lost into excreted acetate. Stronger repression of the acetate consuming Acs in comparison with acetate synthesizing PTA-ACKA together with a decline in TCA cycle flux suggest disruption of acetic acid cycling at the PTA-ACS node (Figure 1). While this node may seem as a futile cycle, the fact is that numerous metabolic tasks involving the intermediate molecules of this cycle-acetyl phosphate (acetyl-P) and acetyl-AMP-are essential for proper *E. coli* growth (Figure 6). For instance, these molecules play a crucial role in bacterial chemotaxis regulation in which flagellar rotation is controlled by the activation level of the response regulator CheY [29] through either phosphotransfer from CheA [30,31] or acetyl-P [31,32], acetylation by acetyl-AMP [33,34] or co-regulation of both mechanisms [29]. It has been also demonstrated that acetyl-P synthesis is vital for EnvZ-independent regulation of outer membrane porins [35], pathogenesis [36] and regulation of several virulence factors [5]. Furthermore, it has been presented that acetyl-P interacts with phosphate concentration regulators PhoB-PhoR [37] and NRI protein which is part of a complex nitrogen sensing system [38]. Acetyl-P is critical for efficient degradation of unfolded or damaged proteins by ATP-dependent

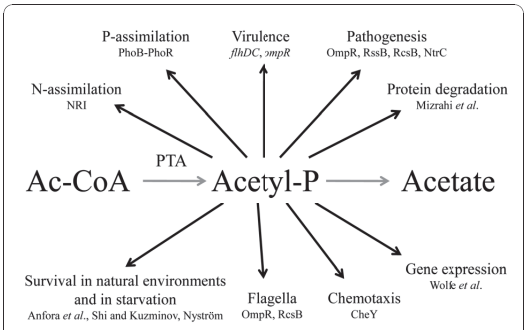

Figure 6 Acetyl-P as an important signal molecule. Ac-CoA, acetyl-CoA; PTA, phosphotransacetylase; refer to text for other abbreviations.

proteases [39]. Altogether, acetyl-P can influence the regulation of almost 100 other genes [40]. Finally, *pta* and/or *ackA* mutations were shown to affect repair-deficient *E. coli* mutants [41] and a *pta* mutant has been reported to be impaired in its ability to survive during glucose starvation, while the *ackA* mutant survived as well as the parent strain [42]. It is important to note that the only known pathway in *E. coli* for acetyl-P synthesis is the PTA-ACKA [5,31]. Taking all the previous into account, we conclude that acetyl-P as well as acetyl-AMP are essential for cellular growth of *E. coli*, and as acetic acid formation is the result of their dephosphorylation, acetic acid should be synthesized and consumed simultaneously during growth to maintain proper balance between metabolites of the PTA-ACS node. This is in agreement with Shin *et al.* [28] who proposed that wild-type *E. coli* constitutively synthesizes acetate even when growing on non-acetogenic carbon source succinate or at low growth rates in carbon limited cultures. It also has to be mentioned that acetic acid is a by-product in the synthesis of cysteine, methionine and arginine, covering around 0.4 mmol/g DCW (Additional file 4). Based on our experimental and literature data, production and re-assimilation of acetate might be over 1 mmol/g DCW at $\mu = 0.2\ h^{-1}$ (Text S2 in Additional file 1) which further supports the hypothesis of the necessity for constant acetic acid synthesis.

A similar regulation for overflow metabolism of acetate was posed for *Saccharomyces cerevisiae* by Postma and co-workers: they postulated that acetate accumulation is the result of insufficient acetyl-CoA synthetase activity for the complete functioning of the pyruvate dehydrogenase bypass because of glucose repression of ACS at high growth rates [43]. The hypothesis proposed here is also consistent with the observation that an *acs* mutant of *E. coli* accumulated acetate in chemostat cultures at dilution rate (D) $0.22\ h^{-1}$ whereas acetate

overflow was started in wild-type at a higher D = 0.35 h^{-1} [28]. Furthermore, it has been shown that over-expression of *acs* [44] and constitutively expressed *acs* together with glyoxylate shunt repressors *iclR* and *fadR* mutant resulted in a significant reduction in acetate accumulation in glucose batch fermentations [28]. Adams and co-workers showed that as a result of micro-evolution, *E. coli* increased acetate consumption capability by over-expressing Acs (not AckA) [45,46], further supporting the connection between Acs activity and acetate accumulation.

As Acs down-regulation is responsible for triggering overflow metabolism and the resulting accumulation of acetate is detrimental to cellular growth, it bears questioning why *E. coli* has not evolved towards maintaining sufficient Acs levels for acetate assimilation in all growth conditions. Growth conditions in *E. coli* native environments are rough as concentrations of utilizable carbon sources including acetate are in the low mg L^{-1} range and access to nutrients is troublesome [47]. These harsh conditions force *E. coli* to make its metabolism ready for scavenging all possible carbon sources including acetate. However, in nutrient rich laboratory conditions, *E. coli* focuses on anthropic growth [48] and biomass production rate, primarily realized by enhancing readily oxidizable substrate (glucose) uptake kinetics which in turn results in Acs repression through CCR and thus, acetate accumulation [46]. This indicates that an active Acs is not essential for rapid growth for *E. coli*. It seems that maintaining high expression levels of acetate assimilation components (and also other alternative substrates ones) is energetically not favorable at higher growth rates. Moreover, as the space on cell membrane is limited and as *E. coli* achieves more rapid growth probably by increasing the number of glucose transport machinery components on the membrane, using area for alternative substrate transport proteins is not beneficial for faster growth. Interestingly, even in one of its natural environments-urinary tract-where a continuous dilution of acetate occurs, it has been shown that metabolizing acetate to acetyl-CoA by Acs is not essential for normal *E. coli* colonization as PTA-ACKA pathway and maintenance of a proper intracellular acetyl-P concentration are necessary for colonizing murine urinary tract [32].

Based on all the points discussed above, PTA-ACS might function as a futile cycle to provide rapid regulation of acetyl-P concentration in the cell for an active chemotaxis that is vital in natural nutrient-depleted environments, fighting against other organisms (acetate production), pathogenesis, biofilm formation etc. This hypothesis is consistent with the fact that the flagellar assembly and regulation operon (*tar-tap-cheRBYZ*) was more intensively expressed at lower growth rates (Additional file 2) where residual glucose concentration is smaller.

Concerning Acs down-regulation, it is possible that CCR is responsible for its repression as proposed by Treves *et al.* [46] showing the link between ACS expression level and acetate accumulation. In our experiments, it was shown that activation of CCR and repression of Acs take place simultaneously prior to the start of overflow metabolism (Figure 1 and Figure 4). As it is well known that CCR is initiated by the presence of glucose in the medium [49,50], we propose that increasing residual glucose concentration accompanying smooth rise of dilution rate in A-stat triggers Acs down-regulation by CCR. The cAMP-Crp complex is one of the major players in CCR of *E. coli* as cAMP binding to Crp drastically increases its affinity towards activating the promotors of catabolic enzymes, including Acs [6,49,50]. We measured a 1.5-fold decrease in Crp expression with increasing growth rate (Figure 1) that is in agreement with the data in the literature [51]. In addition, when *E. coli* mutant defective in the gene *crp* was cultivated in glucose-limited chemostat at a low D = 0.10 h^{-1}, it accumulated acetate whereas the wild-type did not [52]. Furthermore, it exhibited a 34% higher biomass yield relative to the wild-type-this increase might be explained by reduced ATP wasting in the acetate futile cycle, which can be directed to biomass synthesis. Moreover, Khankal *et al.* [53] noted that *E. coli* CRP* mutants that do not require Crp binding to cAMP to activate the expression of catabolic genes showed lowered glucose effect on xylose consumption, 3.6 times higher *acs* expression levels and secreted substantially less acetate in xylitol producing batch fermentations. The connection between cAMP concentration and acetic acid consumption capability, together with the two-phase acetate accumulation profile observed in A-stat and D-stat cultures (Figure 2 and Figure 5) suggests a correlation between increasing residual glucose concentration mediated cAMP-Crp repression and acetate accumulation. Thus, cAMP-Crp dependent regulation of Acs transcribing *acs-yjcH-actP* operon might be a reason for acetate excretion, as also proposed by Veit *et al.* [10]. Our hypothesis of the CCR mediated acetate overflow metabolism is as well in agreement with the fact that rising glucose lowers the intracellular Crp level through the autoregulatory loop of the *crp* gene [54]. However, other mechanisms can also be involved in Acs down-regulation, for example by Cra (Figure 1). Indeed, Sarkar and colleagues have shown that glucose uptake and acetate production rates increased with a decrease of acetate consumption in an *E. coli cra* mutant [55].

What could be the biological relevance of the disruption of the PTA-ACS node? Firstly, decline of the ATP-spending PTA-ACS cycle throughput with increasing growth rate points to possible lower ATP spilling (our model calculations). Secondly, disruption of the PTA-

ACS node decreases the energy needed for expression of this cycle's components. As the disruption of PTA-ACS cycle is CCR-mediated, repression of other alternative substrate transport and utilization enzymes by CCR enables to save additional energy. This could all lead to the decrease of ATP production as was indicated by the diminishing TCA cycle fluxes (Figure 2). Hence, it is plausible that cells repress (by CCR) the expression levels of alternative substrate utilization components (including Acs) for making space on the cell membrane for more preferred substrate (glucose) utilization and ATP producing components to achieve faster growth (see above).

Finally, it was demonstrated that highly reproducible A-stat data are well comparable to chemostat at the level of major growth characteristics and transcriptome, hence quasi steady state data from A-stat can be considered steady state representative (Table 1; Figure S1 and Figure S3 in Additional file 1). Furthermore, as shown also by Postma *et al*. for *S. cerevisiae* [43], chemostat is not fully suitable for characterization of dilution rate dependent metabolic transitions, whereas A-stat should be considered an appropriate tool for this. A-stat is especially well suited for the studies of the details of transient metabolism processes. Dynamic behavior of acetate, cAMP etc. with increasing specific growth rate (Figure 2 Figure 3 and Figure S1 in Additional file 1) and change in acetic acid consumption capability in the two-substrate A-stat (Figure S2 in Additional file 1) could be cited as good examples of the latter.

Conclusion

This study is an excellent example of how a systems biology approach using highly reproducible advanced cultivation methods coupled with multiple 'omics analysis and metabolic modeling allowed to propose a new possible regulation mechanism for overflow metabolism in *E. coli*: acetate overflow is triggered by carbon catabolite repression mediated Acs down-regulation resulting in decreased assimilation of acetate produced by Pta, and disruption of the PTA-ACS node. The practical implications derived from this could lead to better engineering of *E. coli* in overcoming several metabolic obstacles, increasing production yields etc.

Methods

Bacterial strain, medium and continuous cultivation conditions

The *E. coli* K12 MG1655 (λ^- F^- *rph-1Fnr*$^+$; Deutsche Sammlung von Mikroorganismen und Zellkulturen, Germany) strain was used in all experiments. Growth and physiological characteristics in accelerostat (A-stat) cultivations were determined using a defined minimal medium as described before by Nahku *et al*. [51], except 4.5 g/L α-(D)-glucose and 100 μl L^{-1} Antifoam C (Sigma Aldrich, St. Louis, LO) was used. The latter was also used in dilution rate stat (D-stat) experiments as the main cultivation medium. In addition, a second medium was used in D-stat where the main medium was supplemented by acetic acid and prepared as follows: 300 ml medium was withdrawn from the main cultivation medium and supplemented with 3 ml of glacial acetic acid (99.9%). One A-stat experiment (referred to as two-substrate A-stat) was carried out with the same medium as other A-stats, but in addition supplemented with acetic acid (final concentration 5 mM).

The continuous (both A-stat and D-stat) cultivation system consisted of 1.25 L Biobundle bioreactor (Applikon Biotechnology B.V., Schiedam, the Netherlands) controlled by an ADI 1030 biocontroller (Applikon Biotechnology B.V.) and a cultivation control program "BioXpert NT" (Applikon Biotechnology B.V.). The system was equipped with OD, pH, pO_2, CO_2 and temperature sensors. The bioreactor was set on a balance whose output was used as the control variable to ensure constant culture volume (300 ± 1 mL). Similarly, the inflow was controlled through measuring the mass of the fresh culture medium.

A-stat cultivation system and control algorithms used are described in more detail in our previous works [24,51,56]. Dilution rate stat (D-stat) is a continuous cultivation method where dilution rate is constant as in a chemostat while an environmental parameter is smoothly changed [24]. The D-stat experiments in this study were carried out with a slight modification: instead of changing an environmental parameter, two different media were used to keep dilution rate constant. After achieving steady state conditions in chemostat using minimal medium supplemented with glucose, addition of the second medium complemented with glucose and acetic acid was started. The feeding rate of the initial medium was decreased at the same time, resulting in constant glucose concentration in the feed. The acetic acid concentration in the bioreactor as a result of inflow has to be determined to enable precise acetic acid consumption/production rate calculation for the bacteria. Hence, increase of acetic acid concentration in bioreactor was calculated and validated in duplicate non-inoculatedD-stat test experiments producing an average standard deviation of 1.24 mM between calculated and measured acetic acid concentrations.

All continuous cultivation experiments were carried out at 37°C, pH 7 and under aerobic conditions (air flow rate 150 ml min^{-1}) with an agitation speed of 800 rpm. Four A-stat cultivations were performed with acceleration rate (a) 0.01 h^{-2}. Duplicate D-stat experiments were performed at dilution rates 0.10; 0.30; 0.505 ± 0.005 h^{-1} and single experiments at 0.19; 0.24;

0.40; 0.45 h^{-1}. The acetic acid addition profile was set to achieve 32 ± 6 mM and 58 ± 5 mM in 7 hours inside the bioreactor for experiments at dilution rates 0.10-0.24 h^{-1} and 0.30-0.51 h^{-1}, respectively. The growth characteristics of the bacteria were calculated on the basis of total volume of medium pumped out from bioreactor (L), biomass (g DCW), organic acid concentrations in culture medium (mM) and CO_2 concentration in the outflow gas (mM). Formulas were as described in a previous study [24]. It should be noted that the absolute CO_2 concentrations could be error-prone due to measurement difficulties. However, this does not influence the dynamic pattern of specific CO_2 production rate (r_{CO2}) during specific growth rate increase.

Analytical methods

The concentrations of organic acids (lactate, acetate and formate), ethanol and glucose in the culture medium were determined by HPLC and cellular dry weight (expressed as DCW) as described by Nahku *et al*. [51].

Protein expression analysis

Refer to Text S1 in Additional file 1 for detailed description. Shortly, protein expression ratios for around 1600 proteins (identified for each growth rate at a > 95% confidence interval in average from 89,303 distinct 2 or more high-confidence peptides) were generated from mass spectrometric spectra by firstly calculating the ratios between continuous cultivation samples at specific growth rates 0.10 ± 0.01 h^{-1} (chemostat point prior to the start of acceleration in A-stat); 0.20 ± 0.01; 0.26; 0.30 ± 0.01; 0.40 ± 0.00; 0.49 ± 0.01 h^{-1} and batch sample grown on medium containing $^{15}NH_4Cl$ as the only source of ammonia. Secondly, the ratios between the mentioned specific growth rates with chemostat point (μ = 0.10 ± 0.01 h^{-1}) for two biological replicates were calculated to yield protein expression levels for respective specific growth rates. Protein (and gene) expression measurement results are shown in Additional file 2. Proteomic analysis data is also available at the PRIDE database [57]http://www.ebi.ac.uk/pride under accession numbers 12189-12199 (username: review74613, password: Ge9T48e8). The data was converted using PRIDE Converter http://code.google.com/p/pride-converter [58].

Gene expression profiling

DNA microarray analysis of 4,321 transcripts was conducted with the Agilent platform using the data from one A-stat cultivation (a = 0.01 h^{-2}), and gene expression ratios between specific growth rates 0.21; 0.26; 0.31; 0.36; 0.40; 0.48 h^{-1} and μ = 0.11 h^{-1} were calculated. Transcript spot intensities of chemostat sample (sample from D-stat prior to acetic acid addition) from μ = 0.51 h^{-1} and A-stat μ = 0.48 h^{-1} were used for the two method's comparison at transcriptome level. Gene (and protein) expression measurement results are shown in Additional file 2. DNA microarray data is also available at NCBI Gene Expression Omnibus (Reference series: GSE23920). The details of the procedure are provided in Text S1 in Additional file 1.

Metabolome analysis

Sampling was carried out by the rapid centrifugation method. Acquity UPLC (Waters, Milford, MA) together with end-capped HSS C18 T3 1.8 μm, 2.1 × 100 mm column for compound separation coupled to TOF-MS with an electrospray ionization (ESI) source was used for detection (LCT Premiere, Waters). The details of the procedure are provided in Text S1 in Additional file 1.

Additional material

Additional file 1: Detailed Methods (Text S1); calculation of acetate reconsumption (Text S2); Supplementary Figures S1-S5.

Additional file 2: Growth rate dependent gene (one A-stat) and average protein expression changes of two A-stat experiments with *Escherichia coli* K12 MG1655. Transcriptome and proteome analysis results, also with standard deviations.

Additional file 3: Gene spot intensities of A-stat at μ = 0.48 h^{-1} and chemostat at μ = 0.51 h^{-1} experiments with *Escherichia coli* K12 MG1655. Data for A-stat and chemostat transcriptome comparison.

Additional file 4: Simplified metabolic flux analysis. Detailed description of model calculations with simplified metabolic flux analysis.

Acknowledgements

The financial support for this research was provided by the Enterprise Estonia project EU29994, and Ministry of Education, Estonia, through the grant SF0140090s08. The authors would like to thank Lauri Peil and Elina Pelonen for help in carrying out 'omics analysis.

Author details

[1]Tallinn University of Technology, Department of Chemistry, Akadeemia tee 15, 12618 Tallinn, Estonia. [2]Competence Centre of Food and Fermentation Technologies, Akadeemia tee 15b, 12618 Tallinn, Estonia. [3]Tallinn University of Technology, Department of Food Processing, Ehitajate tee 5, 19086 Tallinn, Estonia.

Authors' contributions

KV, KA, and RV drafted the manuscript. RN, PL, and LA helped in preparing the manuscript. KV, RN, and PL designed and performed the experiments. KV analysed the experimental data. RN, PL, and LA carried out the 'omics analysis. KV, KA, and RV guided and coordinated the project. All authors read and approved the manuscript.

Received: 16 June 2010 Accepted: 1 December 2010
Published: 1 December 2010

References

1. Eiteman MA, Altman E: **Overcoming acetate in *Escherichia coli* recombinant protein fermentations.** *Trend Biotechnol* 2006, **24**:530-536.
2. Clomburg JM, Gonzalez R: **Biofuel production in *Escherichia coli*: the role of metabolic engineering and synthetic biology.** *Appl Microbiol Biotechnol* 2010, **86**:419-434.

3. Nakano K, Rischke M, Sato S, Märkl H: **Influence of acetic acid on the growth of *Escherichia coli* K12 during high-cell-density cultivation in a dialysis reactor.** *Appl Microbiol Biotechnol* 1997, **48**:597-601.
4. Contiero J, Beatty CM, Kumari S, DeSanti CL, Strohl WR, Wolfe AJ: **Effects of mutations in acetate metabolism in high-cell-density growth of *Escherichia coli*.** *J Ind Microbiol Biotechnol* 2000, **24**:421-430.
5. Wolfe AJ: **The acetate switch.** *Microbiol Mol Biol Rev* 2005, **69**:12-50.
6. Kumari S, Beatty CM, Browning DF, Busby SJ, Simel EJ, Hovel-Miner G, Wolfe AJ: **Regulation of acetyl coenzyme A synthetase in *Escherichia coli*.** *J Bacteriol* 2000, **182**:4173-4179.
7. Han K, Lim HC, Hong J: **Acetic acid formation in *Escherichia coli* fermentation.** *Biotechnol Bioeng* 1992, **39**:663-671.
8. Farmer WR, Liao JC: **Reduction of aerobic acetate production by *Escherichia coli*.** *Appl Environ Microbiol* 1997, **63**:3205-3210.
9. Majewski RA, Domach MM: **Simple constrained-optimization view of acetate overflow in *E. coli*.** *Biotechnol Bioeng* 1990, **35**:732-738.
10. Veit A, Polen T, Wendisch V: **Global gene expression analysis of glucose overflow metabolism in *Escherichia coli* and reduction of aerobic acetate formation.** *Appl Microbiol Biotechnol* 2007, **74**:406-421.
11. Varma A, Palsson BO: **Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110.** *Appl Environ Microbiol* 1994, **60**:3724-3731.
12. Paalme T, Elken R, Kahru A, Vanatalu K, Vilu R: **The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach.** *Antonie van Leeuwenhoek* 1997, **71**:217-230.
13. Kayser A, Weber J, Hecht V, Rinas U: **Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. I. Growth-rate-dependent metabolic efficiency at steady state.** *Microbiology* 2005, **151**:693-706.
14. El-Mansi M: **Flux to acetate and lactate excretions in industrial fermentations: Physiological and biochemical implications.** *J Ind Microbiol Biotechnol* 2004, **31**:295-300.
15. De Mey M, De Maeseneire S, Soetaert W, Vandamme E: **Minimizing acetate formation in *E. coli* fermentations.** *J Ind Microbiol Biotechnol* 2007, **34**:689-700.
16. El-Mansi EM, Holms WH: **Control of carbon flux to acetate excretion during growth of *Escherichia coli* in batch and continuous cultures.** *J Gen Microbiol* 1989, **135**:2875-2883.
17. Yang Y-T, Bennett GN, San K-Y: **Effect of inactivation of *nuo* and *ackA-pta* on redistribution of metabolic fluxes in *Escherichia coli*.** *Biotechnol Bioeng* 1999, **65**:291-297.
18. Dittrich CR, Vadali RV, Bennett GN, San K-Y: **Redistribution of metabolic fluxes in the central aerobic metabolic pathway of *E. coli* mutant strains with deletion of the *ackA-pta* and *poxB* pathways for the synthesis of isoamyl acetate.** *Biotechnol Prog* 2005, **21**:627-631.
19. Abdel-Hamid AM, Attwood MM, Guest JR: **Pyruvate oxidase contributes to the aerobic growth efficiency of *Escherichia coli*.** *Microbiology* 2001, **147**:1483-1498.
20. Phue J, Noronha SB, Hattacharyya R, Wolfe AJ, Shiloach J: **Glucose metabolism at high density growth of *E. coli* B and *E. coli* K: differences in metabolic pathways are responsible for efficient glucose utilization in *E. coli* B as determined by microarrays and Northern blot analyses.** *Biotechnol Bioeng* 2005, **90**:805-820.
21. Castaño-Cerezo S, Pastor JM, Renilla S, Bernal V, Iborra JL, Cánovas M: **An insight into the role of phosphotransacetylase (*pta*) and the acetate/acetyl-CoA node in *Escherichia coli*.** *Microb Cell Fact* 2009, **8**:54.
22. Zhang W, Li F, Nie L: **Integrating multiple 'omics' analysis for microbial biology: application and methodologies.** *Microbiology* 2010, **156**:287-301.
23. Paalme T, Kahru A, Elken R, Vanatalu K, Tiisma K, Vilu R: **The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate.** *J Microbiol Methods* 1995, **24**:145-153.
24. Kasemets K, Drews M, Nisamedtinov I, Adamberg K, Paalme T: **Modification of A-stat for the characterization of microorganisms.** *J Microbiol Methods* 2003, **55**:187-200.
25. Saier MH, Ramseier TO: **The Catabolite Repressor/Activator (Cra) Protein of Enteric Bacteria.** *J Bacteriol* 1996, **178**:3411-3417.
26. Vemuri GN, Altman E, Sangurdekar DP, Khodursky AB, Eiteman MA: **Overflow metabolism in *Escherichia coli* during steady-state growth: transcriptional regulation and effect of the redox ratio.** *Appl Environ Microbiol* 2006, **72**:3653-3661.
27. Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, Ho PY, 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.
28. Shin S, Chang D, Pan JG: **Acetate Consumption Activity Directly Determines the Level of Acetate Accumulation During *Escherichia coli* W3110 Growth.** *J Microbiol Biotechnol* 2009, **19**:1127-1134.
29. Barak R, Abouhamad WN, Eisenbach M: **Both acetate kinase and acetyl coenzyme A synthetase are involved in acetate-stimulated change in the direction of flagellar rotation in *Escherichia coli*.** *J Bacteriol* 1998, **180**:985-988.
30. Da Re SS, Deville-Bonne D, Tolstykh T, V ron M, Stock JB: **Kinetics of CheY phosphorylation by small molecule phosphodonors.** *FEBS Lett* 1999, **457**:323-326.
31. Mayover TL, Halkides CJ, Stewart RC: **Kinetic characterization of CheY phosphorylation reactions: comparison of P-CheA and small-molecule phosphodonors.** *Biochemistry* 1999, **38**:2259-2271.
32. Klein AH, Shulla A, Reimann Sa, Keating DH, Wolfe AJ: **The intracellular concentration of acetyl phosphate in *Escherichia coli* is sufficient for direct phosphorylation of two-component response regulators.** *J Bacteriol* 2007, **189**:5574-5581.
33. Barak R, Welch M, Yanovsky A, Oosawa K, Eisenbach M: **Acetyladenylate or its derivative acetylates the chemotaxis protein CheY *in vitro* and increases its activity at the flagellar switch.** *Biochemistry* 1992, **31**:10099-10107.
34. Yan J, Barak R, Liarzi O, Shainskaya A, Eisenbach M: ***In vivo* acetylation of CheY, a response regulator in chemotaxis of *Escherichia coli*.** *J Mol Biol* 2008, **376**:1260-1271.
35. Matsubara M, Mizuno T: **EnvZ-independent phosphotransfer signaling pathway of the OmpR-mediated osmoregulatory expression of OmpC and OmpF in *Escherichia coli*.** *Biosci Biotechnol Biochem* 1999, **63**:408-414.
36. Anfora AT, Halladin DK, Haugen BJ, Welch RA: **Uropathogenic *Escherichia coli* CFT073 is adapted to acetatogenic growth but does not require acetate during murine urinary tract infection.** *Infect Immun* 2008, **76**:5760-5767.
37. McCleary W, Stock J: **Acetyl phosphate and the activation of 2-component response regulators.** *J Biol Chem* 1994, **269**:31567-31572.
38. Feng J, Atkinson MR, McCleary W, Stock JB, Wanner BL, Ninfa AJ: **Role of phosphorylated metabolic intermediates in the regulation of glutamine synthetase synthesis in *Escherichia coli*.** *J Bacteriol* 1992, **174**:6061-6070.
39. Mizrahi I, Biran D, Ron EZ: **Involvement of the Pta-AckA pathway in protein folding and aggregation.** *Res Microbiol* 2009, **160**:80-84.
40. Wolfe AJ, Chang D-E, Walker JD, Seitz-Partridge JE, Vidaurri MD, Lange CF, Prüß BM, Henk MC, Larkin JC, Conway T: **Evidence that acetyl phosphate functions as a global signal during biofilm development.** *Mol Microbiol* 2003, **48**:977-988.
41. Shi IY, Kuzminov A: **A Defect in the Acetyl Coenzyme-Acetate Pathway Poisons Recombinational Repair-Deficient Mutants of *Escherichia coli*.** *J Bacteriol* 2005, **187**:1266-1275.
42. Nyström T: **The glucose-starvation stimulon of *Escherichia coli*: induced and repressed synthesis of enzymes of central metabolic pathways and role of acetyl phosphate in gene expression and starvation survival.** *Mol Microbiol* 1994, **12**:833-843.
43. Postma E, Verduyn C, Scheffers Wa, Van Dijken JP: **Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*.** *Appl Environ Microbiol* 1989, **55**:468-477.
44. Lin H, Castro NM, Bennett GN, San K: **Acetyl-CoA synthetase overexpression in *Escherichia coli* demonstrates more efficient acetate assimilation and lower acetate accumulation: a potential tool in metabolic engineering.** *Appl Microbiol Biotechnol* 2006, **71**:870-874.
45. Rosenzweig F, Adams J: **Microbial Evolution in a Simple Unstructured Environment: Genetic Differentiation in *Escherichia coli*.** *Genetics* 1994, **917**:903-917.
46. Treves DS, Manning S, Adams J: **Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*.** *Mol Biol Evol* 1998, **15**:789-797.
47. Franchini AG, Egli T: **Global gene expression in *Escherichia coli* K-12 during short-term and long-term adaptation to glucose-limited continuous culture conditions.** *Microbiology* 2006, **152**:2111-2127.

48. Hardiman T, Lemuth K, Keller M, Reuss M, Siemannherzberg M: **Topology of the global regulatory network of carbon limitation in *Escherichia coli*.** *J Biotechnol* 2007, **132**:359-374.
49. Görke B, Stülke JR: **Carbon catabolite repression in bacteria: many ways to make the most out of nutrients.** *Nat Rev Microbiol* 2008, **6**:613-624.
50. Narang A: **Quantitative effect and regulatory function of cyclic adenosine 5⊠-phosphate in *Escherichia coli*.** *J Biosci* 2009, **34**:445-463.
51. Nahku R, Valgepea K, Lahtvee PJ, Erm S, Abner K, Adamberg K, Vilu R: **Specific growth rate dependent transcriptome profiling of *Escherichia coli* K12 MG1655 in accelerostat cultures.** *J Biotechnol* 2010, **145**:60-65.
52. Nanchen A, Schicker A, Revelles O, Sauer U: **Cyclic AMP-dependent catabolite repression is the dominant control mechanism of metabolic fluxes under glucose limitation in *Escherichia coli*.** *J Bacteriol* 2008, **190**:2323-2330.
53. Khankal R, Chin JW, Ghosh D, Cirino PC: **Transcriptional effects of CRP* expression in *Escherichia coli*.** *J Biol Eng* 2009, **3**:13.
54. Ishizuka H, Hanamura A, Inada T, Aiba H: **Mechanism of the down-regulation of cAMP receptor protein by glucose in *Escherichia coli*: role of autoregulation of the crp gene.** *EMBO J* 1994, **13**:3077-3082.
55. Sarkar D, Siddiquee KA, Araúzo-Bravo MJ, Oba T, Shimizu K: **Effect of *cra* gene knockout together with *edd* and *iclR* genes knockout on the metabolism in *Escherichia coli*.** *Arch Microbiol* 2008, **190**:559-751.
56. Adamberg K, Lahtvee PJ, Valgepea K, Abner K, Vilu R: **Quasi steady state growth of *Lactococcus lactis* in glucose-limited acceleration stat (A-stat) cultures.** *Antonie van Leeuwenhoek* 2009, **95**:219-226.
57. Martens L, Hermjakob H, Jones P, Adamski M, Taylor C, States D, Gevaert K, Vandekerckhove J, Apweiler R: **PRIDE: the proteomics identifications database.** *Proteomics* 2005, **5**:3537-3545.
58. Barsnes H, Vizcaíno JA, Eidhammer I, Martens L: **PRIDE Converter: making proteomics data-sharing easy.** *Nat Biotechnol* 2009, **27**:598-599.

doi:10.1186/1752-0509-4-166
Cite this article as: Valgepea *et al.*: Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase. *BMC Systems Biology* 2010 **4**:166.

PUBLICATION II

Valgepea K, Adamberg K & Vilu R

Decrease of energy spilling in *Escherichia coli* continuous cultures with rising specific growth rate and carbon wasting.

BMC Syst. Biol. 5, 106, (2011)

Valgepea et al. BMC Systems Biology 2011, 5:106
http://www.biomedcentral.com/1752-0509/5/106

BMC Systems Biology

RESEARCH ARTICLE Open Access

Decrease of energy spilling in *Escherichia coli* continuous cultures with rising specific growth rate and carbon wasting

Kaspar Valgepea[1,2], Kaarel Adamberg[2,3] and Raivo Vilu[1,2*]

Abstract

Background: Growth substrates, aerobic/anaerobic conditions, specific growth rate (μ) etc. strongly influence *Escherichia coli* cell physiology in terms of cell size, biomass composition, gene and protein expression. To understand the regulation behind these different phenotype properties, it is useful to know carbon flux patterns in the metabolic network which are generally calculated by metabolic flux analysis (MFA). However, rarely is biomass composition determined and carbon balance carefully measured in the same experiments which could possibly lead to distorted MFA results and questionable conclusions. Therefore, we carried out both detailed carbon balance and biomass composition analysis in the same experiments for more accurate quantitative analysis of metabolism and MFA.

Results: We applied advanced continuous cultivation methods (A-stat and D-stat) to continuously monitor *E. coli* K-12 MG1655 flux and energy metabolism dynamic responses to change of μ and glucose-acetate co-utilisation. Surprisingly, a 36% reduction of ATP spilling was detected with increasing μ and carbon wasting to non-CO_2 by-products under constant biomass yield. The apparent discrepancy between constant biomass yield and decline of ATP spilling could be explained by the rise of carbon wasting from 3 to 11% in the carbon balance which was revealed by the discovered novel excretion profile of *E. coli* pyrimidine pathway intermediates carbamoyl-phosphate, dihydroorotate and orotate. We found that carbon wasting patterns are dependent not only on μ, but also on glucose-acetate co-utilisation capability. Accumulation of these compounds was coupled to the two-phase acetate accumulation profile. Acetate overflow was observed in parallel with the reduction of TCA cycle and glycolysis fluxes, and induction of pentose phosphate pathway.

Conclusions: It can be concluded that acetate metabolism is one of the major regulating factors of central carbon metabolism. More importantly, our model calculations with actual biomass composition and detailed carbon balance analysis in steady state conditions with -omics data comparison demonstrate the importance of a comprehensive systems biology approach for more advanced understanding of metabolism and carbon re-routing mechanisms potentially leading to more successful metabolic engineering.

Background

Escherichia coli exerts a very different gene and protein expression profile under different growth substrates [1], aerobic/anaerobic conditions [2] etc. Specific growth rate (μ) has been shown to be one of the most definite parameters influencing *E. coli* cell physiology as shown by studies of cell size [3,4], biomass composition [5-7], energy metabolism [5,8], transcriptome and proteome [9-11] etc..

To gain insights into the regulation and control mechanisms behind these different phenotype properties, it is useful to know carbon flow patterns in the metabolic network. A widely used tool to calculate quantitative flux values and thereby describe the carbon flow is metabolic flux analysis (MFA). Essentially, MFA calculations need a metabolic network with its stoichiometry, biomass amount and composition, measured steady state carbon influx and outflow-usually as CO_2

* Correspondence: raivo@kbfi.ee
[1]Tallinn University of Technology, Department of Chemistry, Akadeemia tee 15, 12618 Tallinn, Estonia
Full list of author information is available at the end of the article

and by-products. Flux distributions can also be calculated for batch cultures-however, the obtained values have to be considered with great care as the physiological state of cells is constantly changing during growth (*e.g.* μ, by-product production rates). Therefore, MFA is generally carried out with steady state input data from chemostat cultures which provide reproducible and strictly defined physiological state of cells [7,9,12-14].

E. coli mainly uses the consumed carbon for biomass formation and substantial amount of it goes to CO_2 production. The flux (loss) of carbon to CO_2 is closely associated with energy generation (spilling). Carbon usage for biomass synthesis and CO_2 in the carbon balance can be directly measured *in situ* [7,13-15]. However, a notable amount of carbon is lost to many by-products excreted by the cells. The main by-product for most *E. coli* strains in aerobic cultivations is acetic acid [11,13,16]. In addition, accumulation of other compounds such as lactate, formate, pyruvate, ethanol etc. has been observed [7,13,17]. Although excretion of many other compounds besides 'well-known' ones *e.g.* pyrimidine pathway intermediates has been detected [9,18,19], no attention has been drawn on carefully measuring these carbon wasting substances in MFA studies, meaning also that the used metabolic network could be not completely accurate. This can result in a non-closed carbon balance subsequently leading to questionable conclusions. For instance, Taymaz-Nickerel *et al.* accounted a substantial amount of 'leftover carbon' in the carbon balance (7-13%) of *E. coli* continuous cultures to cells lysis which has not been observed before in the literature [7]. Comprehensive carbon balance analysis is, hence, essential for an accurate description of carbon flow and its regulation in the metabolic network under study.

Besides carbon inflow and outflow, biomass composition is another important input parameter in terms of MFA solutions also shown by sensitivity analysis [5]. However, rather than determining it in the same experiments, input values are usually taken from across the literature [7,9,12,15]. Since biomass composition varies under different growth conditions [5-7,15,20] obtaining its values from other studies, with various environmental conditions using various strains of *E. coli*, is another step in addition to a non-comprehensive carbon balance analysis that could possibly lead to distorted MFA results and drawing incorrect conclusions in terms of metabolic regulation. Therefore, we carried out both detailed carbon balance and biomass composition analysis in the same experiments to produce more accurate metabolic flux calculations. What is more, important metabolic switch points and regulation dynamics can be missed when using chemostat cultures. To continuously monitor the flux and energy metabolism dynamic responses to change in μ, we applied the accelerostat (A-stat) [4] continuous cultivation technique which has lately been demonstrated to produce a new regulation mechanism for overflow metabolism of acetate in *E. coli* [11] and an interesting growth efficiency strategy for *Lactococcus lactis* [21]. In addition to A-stat, dilution rate stat (D-stat) method [22] was used to study the effect of glucose-acetate co-utilisation capability on carbon wasting and metabolic flux patterns since this characteristic is proposed to be the key player in *E. coli* overflow metabolism [11].

In short, we detected a 36% reduction of ATP spilling in *E. coli* continuous cultures with increasing μ and carbon wasting under constant biomass yield (Y_{XS}). We propose hypotheses about Y_{XS} maintenance mechanisms and maximal growth limitations for *E. coli* K-12 MG1655. Furthermore, our study revealed novel carbon wasting profiles into pyrimidine pathway intermediates and acetate metabolism governed metabolic flux dynamics that are dependent on μ and glucose-acetate co-utilisation capability.

Results

We carried out three replicate A-stat and four D-stat continuous cultivation experiments at various dilution rates with *E. coli* K-12 MG1655 which growth characteristics are described in detail in Valgepea *et al.* [11]. Carbon balance and biomass composition was carefully determined and the acquired data was used in MFA to obtain better understanding about carbon flow in the metabolic network.

Metabolomic responses to rising μ in A-stat

We detected a two-phase acetate accumulation profile in A-stat which started at $\mu = 0.27 \pm 0.02\ h^{-1}$ (average ± standard deviation) (Figure 1 &2A). Faster accumulation of acetate that took place simultaneously with the abrupt decline in cAMP was witnessed after *E. coli* had reached maximum CO_2 production at $\mu = 0.46 \pm 0.02\ h^{-1}$ (Figure 1). Additionally, we observed considerable excretion of pyrimidine pathway intermediates during increase of μ in three phases (Figure 2A). Dihydroorotate (DHO) and carbamoyl-aspartate (CBASP) accumulated increasingly up to the start of acetate overflow. After overflow switch, DHO started to decline whereas orotate and CBASP levelled off until their levels started to rise again simultaneously (Figure 2A) with the sharp decrease of cAMP and faster accumulation of acetate (Figure 1). Interestingly, in addition to lactate excretion, we observed increased accumulation of a compound once noted in the literature-acetyl-aspartate (NAA)-with rising μ (Figure 2B).

Metabolic flux dynamics with increasing μ

We used data from μ dependent detailed carbon balance and biomass composition analysis to carry out MFA for

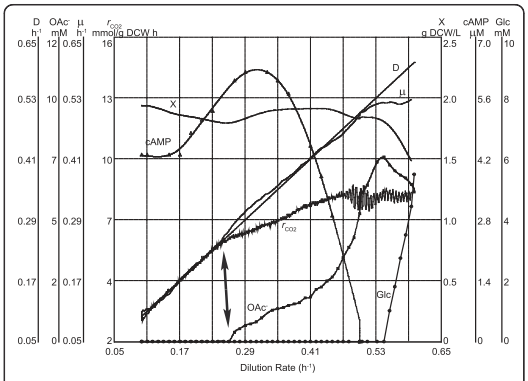

Figure 1 Increasing dilution rate dependent *E. coli* K-12 MG1655 metabolism characterization in one A-stat cultivation. D, dilution rate (h^{-1}); X, biomass concentration (g dry cellular weight (g DCW)/L); μ, specific growth rate (h^{-1}); r_{CO2}, specific CO_2 production rate (mmol/g DCW h); OAc^-, acetate concentration (mM); Glc, glucose concentration (mM); cAMP, cyclic AMP concentration (μM). Arrow indicates the start of overflow metabolism. Start of vertical axes was chosen for better visualization. Reproduced by permission from Valgepea *et al.* [11].

describing the carbon flow in our metabolic network. Biomass composition was dependent on μ (Table S1 in Additional file 1) and its incorporation into MFA calculations was important as shown by up-to 15% difference in flux values compared to using constant biomass composition at different μ (Table S2 in Additional file 1). Our simplified metabolic network (Figure S1 in Additional file 2) consisted of three main pathways-glycolysis, pentose phosphate pathway (PPP), tricarboxylic acid (TCA) cycle-, a part of pyrimidine pathway (to include CBASP, DHO, orotate) and NAA synthesis reaction with 50 fluxes, 22 metabolites taking into account ATP, NADH and NADPH stoichiometry (see Methods and Additional file 2 for details). MFA results for both A-stat and D-stat experiments are given in Additional file 1 (Tables S3-5). Start of acetate overflow triggered reduction of TCA cycle fluxes (Figure 3) also seen by decline of the proportion of CO_2 and NADH produced by TCA cycle (Additional file 3). This subsequently led to induction of PPP fluxes, reduction of glycolysis (Figure 3) and ATP produced from it (Additional file 4). An important central carbon metabolism branch point flux, pyruvate dehydrogenase reaction, reached its maximum throughput at $\mu = 0.42\ h^{-1}$ with concomitant slight increase in glycolysis fluxes resulting in accelerated carbon wasting into by-products (Figure 2 &3). Decrease of Pyk and increase of Ppc and Vprod (excess carbon outflow flux from oxaloacetate in our model) fluxes until acetate accumulation shows that some up-taken carbon was still in excess and excreted as Vprod through Ppc flux (Figure 3). mRNA and protein μ dependent expression dynamics are illustrated as heat maps for all the central carbon metabolism fluxes, see Figure 3A. Correlations between mRNA, protein and flux fold changes and possible metabolic regulation mechanisms will be discussed in a subsequent article. Surprisingly, a strong 32% reduction of ATP spilling (non-growth associated ATP production) was observed after disruption of the PTA-ACS cycle [11] resulting in acetate overflow switch (Figure 4). Reduction of ATP spilling was witnessed since the futile PTA-ACS cycle was included into the model network, see Discussion. Overall ATP production rate increased with μ rise but changed its slope after overflow switch (Additional file 4).

Carbon balance analysis in A-stat

Detailed carbon balance analysis in A-stat showed that most of carbon was used for biomass generation and its proportion relative to CO_2 production increased with rising μ (Figure 5). It can be seen that the carbon balance was not fully closed especially at higher μ which points to loss of carbon into some other not detected compounds. Carbon wasting into by-products increased from 3 to 11% in the carbon balance within the studied μ range (Figure 5 &6). Acetate quickly became the main excreted compound by amount after overflow switch.

Carbon wasting and metabolic flux dynamics in D-stat

In addition to A-stat, we carried out four D-stat experiments at various dilution rates to study the effect of glucose-acetate co-utilisation capability on carbon wasting and metabolic flux patterns. Capability of *E. coli* to co-utilise acetate simultaneously with glucose was strongly repressed with increasing dilution rates (Figure 7). We observed that carbon wasting patterns into orotate, DHO, CBASP, NAA and lactate changed under different co-utilisation properties (Table S5 in Additional file 1). It is interesting to note that the percentage of overall carbon wasting (sum of orotate, DHO, CBASP, NAA and lactate) in the carbon balance was similar (*ca* 5.5%) at the studied dilution rates under very different maximal glucose-acetate co-utilisation capability values (Figure 7). As expected, MFA calculations with D-stat data revealed that acetate consumption fluxes together with glyoxylate shunt and gluconeogenesis were higher under higher glucose-acetate co-utilisation values (Table S5 in Additional file 1). Concomitantly, glycolysis fluxes were down-regulated at lower co-utilisation conditions similarly with pyruvate dehydrogenase.

A-stat comparison with chemostat

A-stats have been shown to produce quantitatively comparable results with chemostats at the level of *E. coli* and *Lactococcus lactis* major growth characteristics (*e.g.*

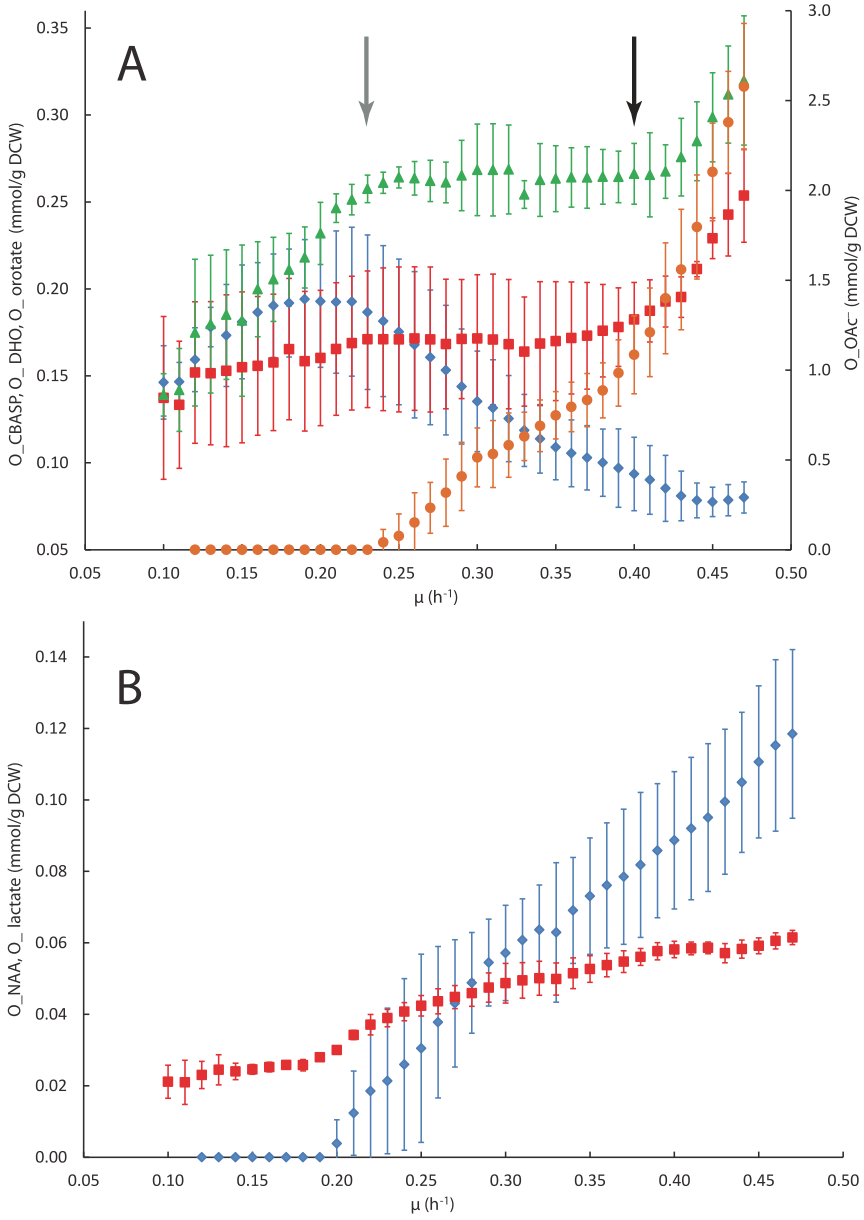

Figure 2 Specific growth rate dependent dynamic carbon wasting profiles in three *E. coli* K-12 MG1655 A-stat cultivations. μ, specific growth rate (h^{-1}). Error bars represent standard deviation of triplicate A-stat experiments. A. Production of compounds per biomass (mmol/g DCW): CBASP, carbamoyl-aspartate (green triangle); DHO, dihydroorotate (blue diamond); orotate (red square); OAc^-, acetate (orange circle). Grey arrow denotes acetate overflow switch with concomitant stop of DHO and CBASP increase whereas black arrow depicts faster acetate accumulation coupled induction of orotate and CBASP excretion. B. Production of compounds per biomass (mmol/g DCW): NAA, acetyl-aspartate (blue diamond); lactate (red square).

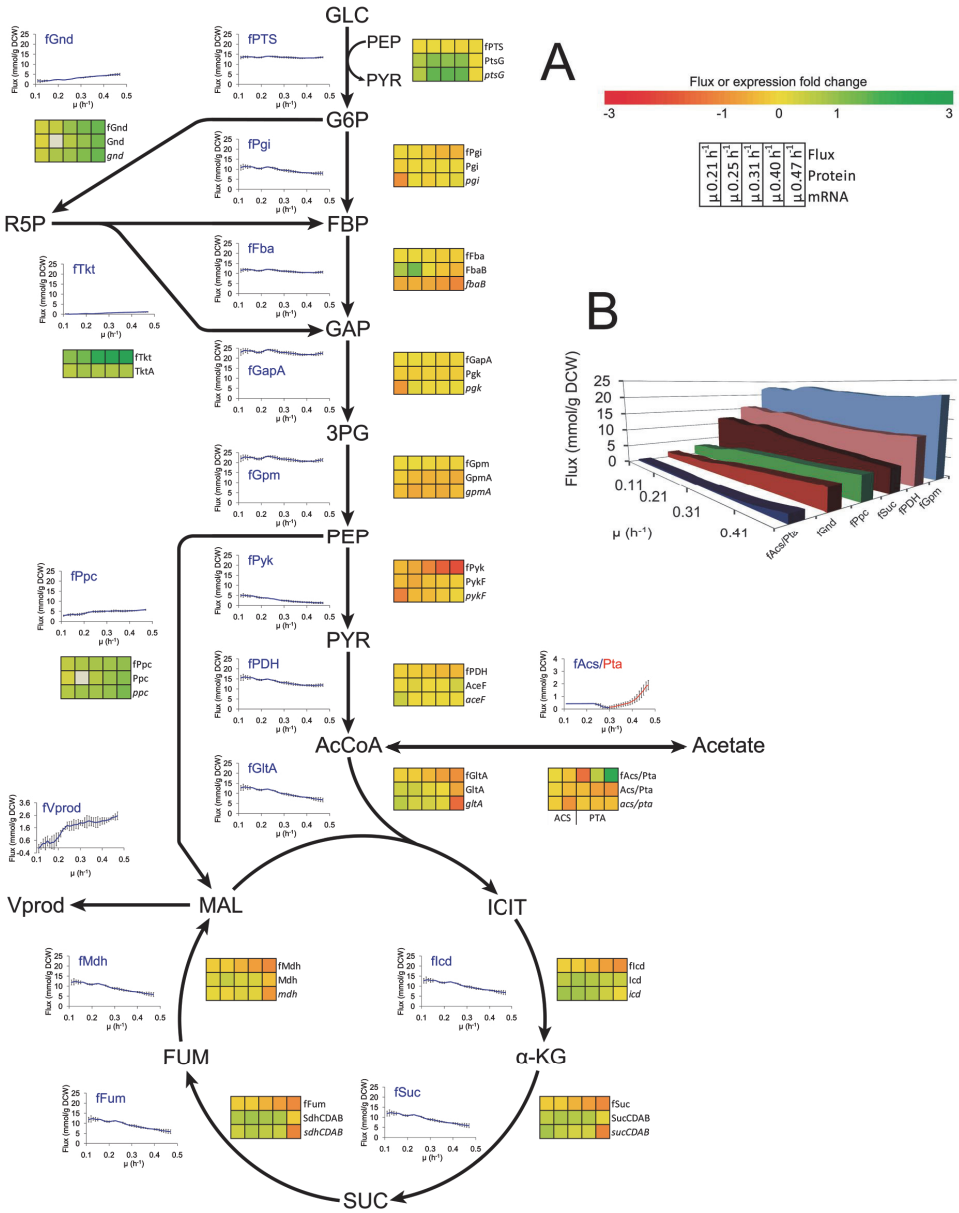

Figure 3 ***E. coli*** **K-12 MG1655 central carbon metabolism flux, protein and mRNA dynamics with rising specific growth rate in three A-stat cultivations**. μ, specific growth rate (h^{-1}). A. Flux or expression fold change (log_2) is calculated for each μ compared to μ = 0.10 h^{-1}; grey box depicts missing value; Acs flux was switched to Pta after μ = 0.31 h^{-1} in MFA since acetate excretion exceeds its production accompanying biosynthesis. Protein and mRNA data taken from Valgepea *et al.* [11]. Error bars represent standard deviation of triplicate A-stat experiments. B. Selected glycolysis, TCA cycle, PPP, gluconeogenesis and acetate related fluxes.

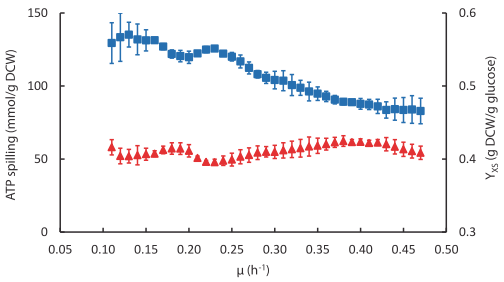

Figure 4 Specific growth rate dependent ATP spilling and biomass yield in three *E. coli* K-12 MG1655 A-stat cultivations. μ, specific growth rate (h^{-1}); ATP spilling, non-growth associated ATP production (blue square); Y_{XS}, biomass yield as gram dry cellular weight per gram consumed glucose (red triangle). Error bars represent standard deviation of triplicate A-stat experiments.

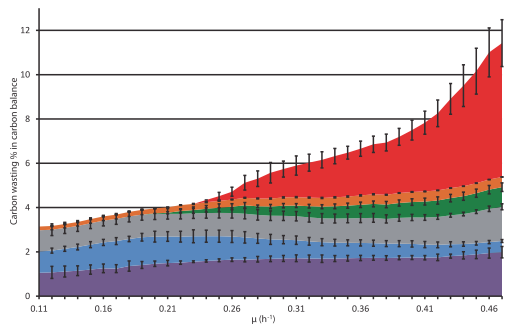

Figure 6 Specific growth rate dependent dynamic carbon wasting profiles in three *E. coli* K-12 MG1655 A-stat cultivations. μ, specific growth rate (h^{-1}). Carbon wasting % in carbon balance into: acetate (red); lactate (orange); NAA, acetyl-aspartate (green); orotate (grey); DHO, dihydroorotate (blue); CBASP, carbamoyl-aspartate (violet). Error bars represent standard deviation of triplicate A-stat experiments.

Y_{XS}, by-product production per biomass) [11,21] and *E. coli* transcriptome [11]. Our detailed carbon balance analysis further confirmed that A-stat and chemostat data are quantitatively comparable (Table 1) which enables to use quasi steady state data from A-stat for steady state modeling calculations.

Discussion

We applied a systems biology approach to study *E. coli* metabolic flux dynamics and possible growth limiting factors. Detailed carbon balance and biomass composition analysis were carried out in A-stat and D-stat cultures to examine the dynamic responses of metabolic fluxes and energy metabolism to change of μ and glucose-acetate co-utilisation capability. A simplified MFA was conducted to map the carbon flow through central carbon metabolism.

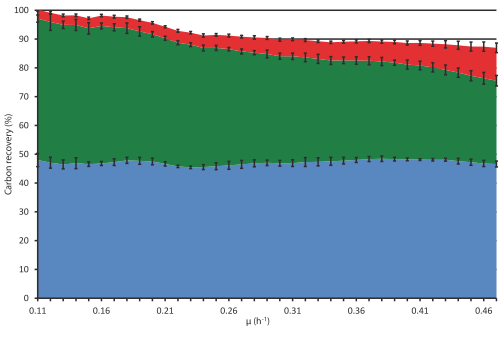

Figure 5 Specific growth rate dependent carbon balance in three *E. coli* K-12 MG1655 A-stat cultivations. μ, specific growth rate (h^{-1}). Carbon recovery % in carbon balance into: biomass (blue); CO_2 (green); sum of carbon wasting into acetate, CBASP, DHO, lactate, NAA, orotate (red). Error bars represent standard deviation of triplicate A-stat experiments.

We detected a two-phase acetate accumulation profile in A-stat which started at μ = 0.27 ± 0.02 h^{-1} (Figure 1). After linear increase, acetate probably starts to accumulate exponentially because of total repression of acetate consuming enzyme, acetyl-CoA synthetase, by carbon catabolite repression [11]. It became clear from MFA calculations that acetate excretion plays an important role in overall flux patterns and ATP metabolism. Firstly, start of acetate excretion reduces carbon flow from the PTA-ACS cycle to acetyl-CoA and central metabolism triggering reduction of TCA cycle fluxes (Figure 3) that can be also seen by decline of the proportion of CO_2 and NADH produced by the TCA cycle (Additional file 3). This was paralleled with induction of PPP, possibly for NADPH regeneration, and reduction of glycolysis (Figure 3). These shifts have been observed in chemostat cultures as well [9,12-14,23]. Secondly, disruption of the PTA-ACS node resulting in acetate overflow strongly reduced ATP spilling (non-growth associated ATP production) which declined 32% with increasing of μ (Figure 4). What is more, slope of the overall ATP production rate changed right after overflow switch (Additional file 4). It can be concluded from the latter that acetate metabolism is one of the major regulating factors of central carbon metabolism which is in good agreement with abundance of literature data.

Decrease in ATP spilling (32%) after overflow switch in A-stat was shown by MFA calculations (Figure 4) and decline of the ATP-spending PTA-ACS node throughput by acetate excretion (Figure 2A). This response in energy metabolism was detected in this study since the futile

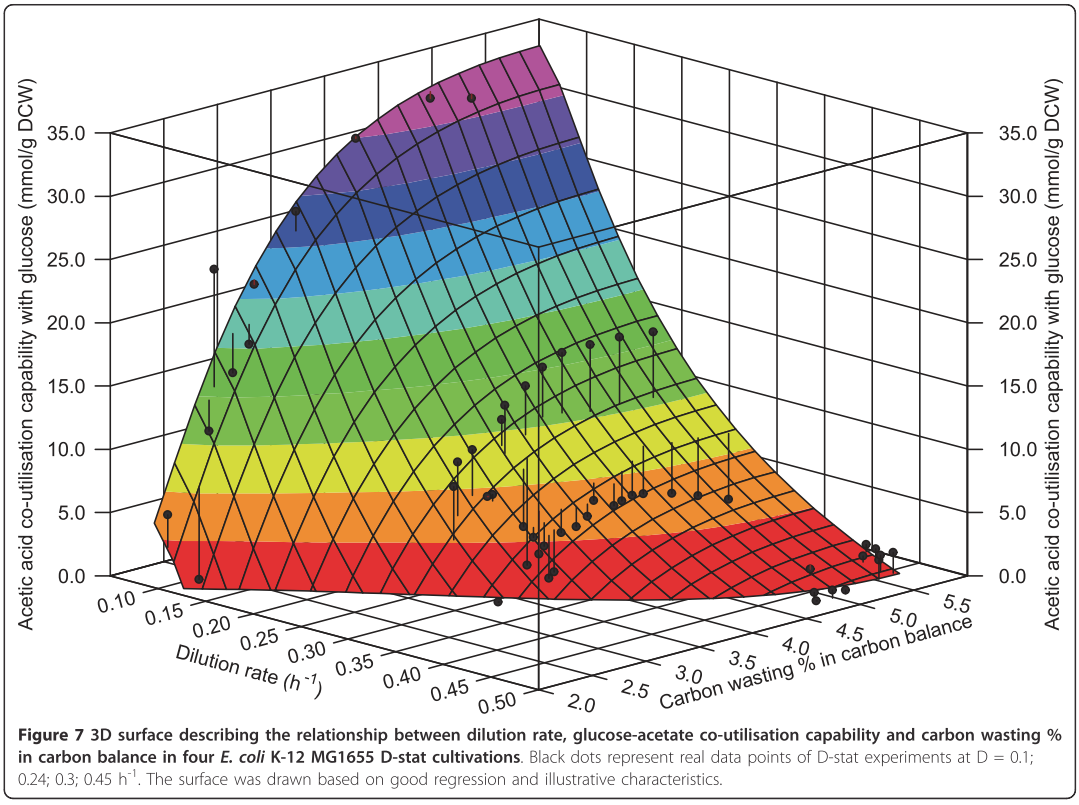

Figure 7 3D surface describing the relationship between dilution rate, glucose-acetate co-utilisation capability and carbon wasting % in carbon balance in four *E. coli* K-12 MG1655 D-stat cultivations. Black dots represent real data points of D-stat experiments at D = 0.1; 0.24; 0.3; 0.45 h^{-1}. The surface was drawn based on good regression and illustrative characteristics.

PTA-ACS cycle was included into the model network. Additionally, change in ATP production rate was also seen with increasing μ (Additional file 4). We have to note that the possibility of our ATP calculations being distorted due to carbon imbalance at higher specific growth rates (Figure 5) cannot be excluded which could lead to underestimation of ATP production. However, this seems rather unlikely since CO_2 measurement precision is constant for all μ. Furthermore, theoretically no other pathway besides the TCA cycle cannot by far produce enough energy under these carbon imbalance conditions so that decrease of ATP spilling would not be observed. We additionally have to point out that our calculations could also overestimate ATP production since a theoretical ratio for ATP generation efficiency in oxidative phosphorylation (P/O = 2) was chosen which can be higher than the *in vivo* value. In any case, the P/O ratio *per se* does not affect the main conclusions of the manuscript since the ratio between ATP spilling and its overall production is independent from the P/O ratio value.

Decrease in ATP spilling (40 mmol/g dry cellular weight (DCW)) might indicate increase of Y_{XS}, however,

Table 1 A-stat and chemostat non-CO_2 by-product production comparison

	μ = 0.10 h^{-1}		μ = 0.24 h^{-1}		μ = 0.30 h^{-1}		μ = 0.45 h^{-1}	
	Chemostat	**A-stat**	**Chemostat**	**A-stat**	**Chemostat**	**A-stat**	**Chemostat**	**A-stat**
O_{CBASP}	0.188	0.139 ± 0.047	0.089	0.218 ± 0.018	0.195	0.268 ± 0.026	0.188	0.299 ± 0.026
O_{DHO}	0.131	0.146 ± 0.021	0.117	0.181 ± 0.043	0.091	0.135 ± 0.029	0.065	0.078 ± 0.008
$O_{lactate}$	ND	0.021 ± 0.005	0.034	0.041 ± 0.003	0.034	0.049 ± 0.006	0.046	0.059 ± 0.002
O_{NAA}	ND	ND	0.006	ND	0.034	0.057 ± 0.013	0.069	0.111 ± 0.021
$O_{orotate}$	0.091	0.137 ± 0.047	0.193	0.171 ± 0.041	0.115	0.171 ± 0.037	0.344	0.229 ± 0.012

Unit is mmol/g DCW. A-stat values represent the average from three independent experiments and standard deviation follows the ± sign. Chemostat values from one experiment. Refer to text for abbreviations. ND, not detected.

it remained constant in our experiments (Figure 4). This apparent discrepancy between the decrease in ATP spilling and constant Y_{XS} (Figure 4) could be explained by the fact that carbon wasting increases from 3 to 11% with rising μ (Figure 6) as follows. As the acetate synthesis/assimilation PTA-ACS is a futile cycle, an equivalent amount of ATP to acetate is concomitantly wasted with production and re-consumption of acetate. Therefore, accumulation of acetate likely triggers a 32% decline of ATP spilling (Figure 4) since re-consumption of acetate (wasting 1 molecule of ATP) decreases with rising μ after overflow switch. This energy save is, however, counteracted by the increase of carbon wasting in the carbon balance from 3 to 11% which results in a constant Y_{XS}. However, *E. coli* might possess additional mechanisms to maintain a constant Y_{XS} under increasing carbon wasting conditions during μ increase.

In addition to metabolic flux dynamics, we described novel carbon wasting profiles in *E. coli* K-12 MG1655 into pyrimidine pathway intermediates orotate, DHO, CBASP, and NAA with rising μ (Figure 2) and under various glucose-acetate co-utilisation capabilities (Table S5 in Additional file 1). Excretion of orotate [18,19], DHO [19], CBASP [19] and NAA [9] by *E. coli* has been noted before. Accumulation of the pyrimidine pathway compounds-orotate, DHO and CBASP-can be explained by the *E. coli* K-12 MG1655 genotype. This specific strain is prone to pyrimidine starvation due to a *rph* frameshift mutation leading to low *pyrE* (encodes PyrE protein which catalyses orotate conversion into orotidine-5-phosphate) expression [24] which could possibly lead to accumulation of precursor molecules which all the latterly mentioned compounds are (Additional file 5). Excretion of a considerable amount of CBASP, DHO, orotate and NAA besides acetate shows that overflow metabolism actually consists of more products than acetate, as generally believed. Detailed by-product measurements enabled us to precisely detect carbon outflow routes for MFA calculations which usually are taken into account predictively either from pyruvate, oxaloacetate, α-ketoglutarate or other potential precursors. For instance, if these product outflows will be excluded from MFA and carbon wasting predicted as pyruvate outflow, pyruvate dehydrogenase, TCA cycle and PEP carboxylase fluxes would deviate by 11, 24 and 60%, respectively, at $\mu = 0.47\ h^{-1}$ from the values calculated by our model (Table S6 in Additional file 1). Although the absolute amount of these excreted substances in the carbon balance is not substantial (less than 5%), linking their accumulation dynamics to μ (or metabolic routes) is relevant for acknowledging the potential imbalance between pyrimidine metabolism, carbon re-consumption and ATP spilling.

DHO and CBASP accumulated increasingly up to the start of acetate overflow (Figure 2A). After overflow switch, DHO started to decline whereas orotate and CBASP levelled off suggesting that carbon flow from the PTA-ACS cycle to acetyl-CoA and central metabolism declines indicated by carbon loss to excreted acetate and decreasing TCA cycle flux patterns (Figure 2A &3). Finally, orotate and CBASP levels started to increase again simultaneously (Figure 2A) with the sharp decrease of cAMP and faster accumulation of acetate (Figure 1). Firstly, this rise could be explained by the high demand for RNA synthesis at higher μ which leads to precursor molecule accumulation because of the low *pyrE* expression. On the other hand, pyruvate dehydrogenase flux reached its maximum at $\mu = 0.42\ h^{-1}$ with concomitant slight increase in glycolysis fluxes that subsequently resulted in accelerated carbon wasting (Figure 3). These observations demonstrate a strong link between overflow metabolism of acetate and carbon wasting into other products.

We validated and quantified, for the first time to our knowledge, excretion of NAA which levels constantly increased with μ in A-stat experiments (Figure 2B). Neither has NAA yet been registered in EcoCyc Ver 15.0 [25] nor there exists a protein catalysing its synthesis for *E. coli* K-12 MG1655 in KEGG Release 58.0 [26]. Based on homology analysis with the available aspartate N-acetyltransferase protein sequences (human and mouse), we propose that a predicted acetyltransferase YjgM could catalyse the formation of NAA in *E. coli* (Additional file 6). This is supported by the fact that YjgM is expressed within the studied μ range [11]. We hypothesise that since oxaloacetate is over-produced in the TCA cycle and this excess carbon cannot be shunted towards PPP because of the discovered 'CBASP-DHO-orotate' bottleneck, the excess carbon is excreted as NAA.

To our knowledge, dilution rate dependent carbon wasting profiles in terms of two-substrate (glucose and acetic acid) co-utilisation has not been studied before. We found that carbon wasting patterns are dependent not only on μ as shown by A-stat, but also on glucose-acetate co-utilisation capability (Figure 7). This finding could be useful for bioprocess development since mixed-substrate growth is commonly used there and loss of carbon is unwanted. What is more, overall carbon wasting into the carbon wasting products in the carbon balance was similar (*ca* 5.5%) under very different maximal glucose-acetate co-utilisation capability values (Figure 7). The latter implies that the quantitative carbon flow through the carbon wasting pathways does not depend on the consumption of additional (to glucose) carbon sources, such as acetate in our case.

Overall carbon wasting in the carbon balance increased with rising μ (Figure 5 &6). Carbon wasting to other substances than acetate *e.g.* orotate, DHO, CBASP, NAA may be caused by an imbalance between the supply of precursors, intermediates of central metabolic network and insufficient use of them for the synthesis of biomass monomers and macromolecules which could be the result of energy limitations. Levels of metabolites from the upper part of energy generating glycolysis, fructose-1,6-bisphosphate and glyceraldehyde-3-phosphate, and TCA cycle components increased with the rise in μ [11], indicating no limits in energy supply at these steps. However, there might be competing futile cycles wasting ATP, and leading to energy limitations, for instance in glycolysis [27] and the PTA-ACS cycle proposed by us [11]. In addition to the pyrimidine synthesis pathway compounds-orotate, DHO and CBASP-, we observed TCA cycle intermediates α-ketoglutarate and isocitrate accumulation, and decline in intracellular ribose-5-phosphate and erythrose-4-phosphate metabolite levels, which are also intermediates of pyrimidine synthesis [11]. This might result in growth limitation by RNA synthesis and ribosome assembly, especially at higher μ. The latter is supported by the fact about the genotype of *E. coli* K-12 MG1655 (*rph* frameshift mutation leading to low *pyrE* expression) which leads to pyrimidine starvation as described above. The latter proposal is in agreement with RNA concentration measurements which showed that RNA amount increased together with μ until 0.40 h^{-1} after which it levelled off (Table S1 in Additional file 1). The possible limitation of maximal μ by RNA synthesis and carbon wasting due to the *E. coli* K-12 MG1655 genotype revealed by this study proposes a way how to increase maximal μ and Y_{XS} which is relevant for the biotechnology industry. Furthermore, all the data referred above show that the details of the regulatory mechanisms of cellular growth need further studying using comprehensive systems biology approaches.

Conclusions

In this work, we described a novel carbon wasting strategy to pyrimidine pathway intermediates and acetate metabolism governed metabolic flux dynamics in *E. coli* continuous cultures dependent on μ and glucose-acetate consumption capability. More importantly, MFA calculations with actual biomass composition revealed a 36% reduction of ATP spilling with increasing μ and carbon wasting to non-CO_2 by-products under constant Y_{XS}. Our model calculations with actual biomass composition and detailed carbon balance analysis in steady state conditions with -omics data comparison demonstrate the importance of a comprehensive systems biology approach for more advanced understanding of carbon re-routing mechanisms.

Methods

Bacterial strain, medium and continuous cultivation conditions

E. coli K-12 MG1655 (λ-F-*rph-1Fnr+*; Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), DSM No.18039) was used in A-stat and D-stat experiments under the following conditions: temperature 37° C, pH 7, agitation speed of 800 rpm, and aerobic conditions (air flow rate 150 ml/min). Three A-stat cultivations were performed with acceleration rate 0.01 h^{-2} and four D-stat experiments at dilution rates 0.10; 0.24; 0.30; 0.45 h^{-1}. A detailed description of medium and cultivation conditions in these experiments has been reported previously [11]. In short, defined minimal medium with 4.5 g/L α-(D)-glucose was used in A-stat experiments. The latter was also used in D-stat experiments as the main cultivation medium, additionally the main medium was supplemented with acetic acid and prepared as follows: 300 ml medium was withdrawn from the main cultivation medium and supplemented with 3 ml of glacial acetic acid (99.9%).

Analytical methods

Biomass concentration was determined gravimetrically as dry cellular weight (expressed as DCW) described by Nahku *et al.* [10].

Samples of culture medium for extracellular metabolome analysis were centrifuged at 14,000 × *g* for 5 min, supernatant were collected and analysed immediately or stored at -20°C until analysis. Glucose and organic acids were analysed by HPLC (Alliance2795 system, Waters, Milford, MA) using a BioRad HPX-87H Aminex ion-exclusion column connected to RI and UV detectors (35°C, flow rate 0.6 ml/min). The column was eluted with 4.1 mM sulphuric acid for glucose, CBASP, lactate, orotate and with 26.5 mM formic acid for acetate, DHO and NAA analysis.

Metabolic flux analysis (MFA)

Simplified metabolic network of *E. coli* K-12 MG1655 for MFA was reconstructed taking into account main metabolic pathways-glycolysis, PPP, TCA cycle-, one pathway from pyrimidine metabolism (to include CBASP, DHO, orotate) and NAA synthesis reaction (Figure S1 in Additional file 2). Our network involved only fluxes (reactions) between branching points (metabolites) whereas linear pathway chains were lumped together. Fully determined and calculable stoichiometric matrix consisted of 22 metabolites and 50 fluxes (24 dependent fluxes, one measured inflow, seven outflow fluxes and 18 calculated fluxes based on biomass composition and stoichiometries of biosynthetic pathways). Biomass composition was shown to be dependent on μ (Table S1 in Additional file 1). Cofactors ATP, NADPH and NADH were considered

in calculations. Refer to Additional file 2 for detailed description for reaction stoichiometries and calculation of biomass composition dependent fluxes. MFA results for both A-stat and D-stat experiments are given in Additional file 1 (Table S3-5).

Additional material

Additional file 1: Specific growth rate dependent biomass composition and metabolic flux analysis results of triplicate *E. coli* K-12 MG1655 A-stat experiments. Specific growth rate dependent *E. coli* K-12 MG1655 biomass monomer composition (Table S1); Effect of specific growth rate dependent biomass composition on MFA results (Table S2); MFA results for triplicate *E. coli* K-12 MG1655 A-stat experiments (Table S3); MFA average results and standard deviations of triplicate *E. coli* K-12 MG1655 A-stat experiments (Table S4); MFA results of four *E. coli* K-12 MG1655 D-stat experiments (Table S5); Effect of taking novel carbon wasting routes into account in *E. coli* K-12 MG1655 A-stat experiments' MFA (Table S6).

Additional file 2: Metabolic flux analysis. Detailed description of model calculations with simplified metabolic flux analysis; Simplified metabolic network scheme of *E. coli* K-12 MG1655 (Figure S1).

Additional file 3: *E. coli* K-12 MG1655 proportion of CO_2 and NADH production by TCA cycle with rising specific growth rate in three A-stat cultivations. μ, specific growth rate (h^{-1}). CO_2 production (blue square); NADH production (red triangle). Error bars represent standard deviation of triplicate A-stat experiments.

Additional file 4: Specific growth rate dependent overall ATP production rate and proportion of ATP production by glycolysis in three *E. coli* K-12 MG1655 A-stat cultivations. μ, specific growth rate (h^{-1}). Overall ATP production (blue squares); ATP production by glycolysis (red triangle). Error bars represent standard deviation of triplicate A-stat experiments.

Additional file 5: *E. coli* K-12 MG1655 pyrimidine pathway *rph* frameshift mutation triggered accumulating precursor compounds. Carbamoyl-P, carbamoyl-phosphate; CBASP, carbamoyl-aspartate; DHO, dihydroorotate; Oro-5P, orotidine-5-phosphate; TCA cycle, tricarboxylic acid cycle; PPP, pentose phosphate pathway; *pyrB*, aspartate carbamoyltransferase; *pyrC*, dihydro-orotase; *pyrD*, dihydro-orotate oxidase; *pyrE*, orotate phosphoribosyltransferase. Gene names are italicised.

Additional file 6: Homology analysis for *E. coli* acetyltransferase prediction. BLAST results for homology analysis with *Homo sapiens* and *Mus musculus* aspartate N-acetyltransferases.

Acknowledgements
The authors thank Ranno Nahku for critical reading of the manuscript. The financial support for this research was provided by the Enterprise Estonia project EU29994, Ministry of Education, Estonia, through the grant SF0140090s08 and Estonian Science Foundation through the grant G8165.

Author details
[1]Tallinn University of Technology, Department of Chemistry, Akadeemia tee 15, 12618 Tallinn, Estonia. [2]Competence Centre of Food and Fermentation Technologies, Akadeemia tee 15b, 12618 Tallinn, Estonia. [3]Tallinn University of Technology, Department of Food Processing, Ehitajate tee 5, 19086 Tallinn, Estonia.

Authors' contributions
KV, KA, and RV designed, guided and coordinated the project. KV performed the experiments, carried out analytics and data analysis. KV and KA were responsible for model calculations. KV drafted the manuscript. KA and RV helped in drafting the manuscript. All authors read and approved the manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 27 April 2011 Accepted: 5 July 2011 Published: 5 July 2011

References

1. Hua Q, Yang C, Oshima T, Mori H, Shimizu K: **Analysis of gene expression in *Escherichia coli* in response to changes of growth-limiting nutrient in chemostat cultures.** *Appl Environ Microbiol* 2004, **70**:2354-2366.
2. Partridge JD, Sanguinetti G, Dibden DP, Roberts RE, Poole RK, Green J: **Transition of *Escherichia coli* from aerobic to micro-aerobic conditions involves fast and slow reacting regulatory components.** *J Biol Chem* 2007, **282**:11230-11237.
3. Mongold JA, Lenski RE: **Experimental rejection of a nonadaptive explanation for increased cell size in *Escherichia coli*.** *J Bacteriol* 1996, **178**:5333-5334.
4. Paalme T, Kahru A, Elken R, Vanatalu K, Tiisma K, Vilu R: **The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate (A-stat).** *J Microbiol Methods* 1995, **24**:145-153.
5. Pramanik J, Keasling JD: **Stoichiometric model of *Escherichia coli* metabolism: incorporation of growth-rate dependent biomass composition and mechanistic energy requirements.** *Biotechnol Bioeng* 1997, **56**:398-421.
6. Špitsmeister M, Adamberg K, Vilu R: **UPLC/MS based method for quantitative determination of fatty acid composition in Gram-negative and Gram-positive bacteria.** *J Microbiol Methods* 2010, **82**:288-295.
7. Taymaz-Nikerel H, Borujeni AE, Verheijen PJT, Heijnen JJ, van Gulik WM: **Genome-derived minimal metabolic models for *Escherichia coli* MG1655 with estimated in vivo respiratory ATP stoichiometry.** *Biotechnol Bioeng* 2010, **107**:369-381.
8. Fuhrer T, Sauer U: **Different biochemical mechanisms ensure network-wide balancing of reducing equivalents in microbial metabolism.** *J Bacteriol* 2009, **191**:2112-2121.
9. Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, Ho PY, 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.
10. Nahku R, Valgepea K, Lahtvee PJ, Erm S, Abner K, Adamberg K, Vilu R: **Specific growth rate dependent transcriptome profiling of *Escherichia coli* K12 MG1655 in accelerostat cultures.** *J Biotechnol* 2010, **145**:60-65.
11. Valgepea K, Adamberg K, Nahku R, Lahtvee PJ, Arike L, Vilu R: **Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase.** *BMC Syst Biol* 2010, **4**:166.
12. Nanchen A, Schicker A, Sauer U: **Nonlinear dependency of intracellular fluxes on growth rate in miniaturized continuous cultures of *Escherichia coli*.** *Appl Environ Microbiol* 2006, **72**:1164-1172.
13. Kayser A, Weber J, Hecht V, Rinas U: **Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. I. Growth-rate-dependent metabolic efficiency at steady state.** *Microbiology* 2005, **151**:693-706.
14. Zhao J, Shimizu K: **Metabolic flux analysis of *Escherichia coli* K12 grown on ^{13}C-labeled acetate and glucose using GC-MS and powerful flux calculation method.** *J Biotechnol* 2003, **101**:101-117.
15. Emmerling M, Dauner M, Ponti A, Fiaux J, Hochuli M, Szyperski T, Wüthrich K, Bailey JE, Sauer U: **Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*.** *J Bacteriol* 2002, **184**:152-164.
16. El-Mansi M, Holms WH: **Control of carbon flux to acetate excretion during growth of *Escherichia coli* in batch and continuous cultures.** *J Gen Microbiol* 1989, **135**:2875-2883.
17. Chen R, Yap WM, Postma PW, Bailey JE: **Comparative studies of *Escherichia coli* strains using different glucose uptake systems: Metabolism and energetics.** *Biotechnol Bioeng* 1997, **56**:583-590.
18. Womack JE, O Donovan GA: **Orotic acid excretion in some wild-type strains of *Escherichia coli* K-12.** *J Bacteriol* 1978, **136**:825-827.
19. Yates RA, Pardee AB: **Control of pyrimidine biosynthesis in *Escherichia coli* by a feed-back mechanism.** *J Biol Chem* 1956, **221**:757-770.
20. Pramanik J, Keasling JD: **Effect of *Escherichia coli* biomass composition on central metabolic fluxes predicted by a stoichiometric model.** *Biotechnol Bioeng* 1998, **60**:230-238.

21. Lahtvee PJ, Adamberg K, Arike L, Nahku R, Aller K, Vilu R: **Multi-omics approach to study the growth efficiency and amino acid metabolism in *Lactococcus lactis* at various specific growth rates.** *Microb Cell Fact* 2011, **10**:12.
22. Kasemets K, Drews M, Nisamedtinov I, Paalme T, Adamberg K: **Modification of A-stat for the characterization of microorganisms.** *J Microbiol Methods* 2003, **55**:187-200.
23. Vazquez A, Beg QK, Demenezes MA, Ernst J, Bar-Joseph Z, Barabási AL, Boros LG, Oltvai ZN: **Impact of the solvent capacity constraint on *E. coli* metabolism.** *BMC Syst Biol* 2008, **2**:7.
24. Jensen KF: **The *Escherichia coli* K-12 "wild types" W3110 and MG1655 have an *rph* frameshift mutation that leads to pyrimidine starvation due to low *pyrE* expression levels.** *J Bacteriol* 1993, **175**:3401-3407.
25. Collado-Vides J, Santos-Zavaleta A, Peralta-Gil M, Gama-Castro S, Muniz-Rascado L, Bonavides-Martinez C, Paley S, Krummenacker M, Altman T, Kaipa P, Spaulding A, Pacheco J, Latendresse M, Fulcher C, Sarker M, Shearer AG, Mackie A, Paulsen I, Karp PD: **EcoCyc: a comprehensive database of *Escherichia coli* biology.** *Nucleic Acids Res* 2010, **39**:D583-590.
26. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M: **KEGG for representation and analysis of molecular networks involving diseases and drugs.** *Nucleic Acids Res* 2010, **38**:D355-360.
27. Russell JB: **The energy spilling reactions of bacteria and other organisms.** *J Mol Microbiol Biotechnol* 2007, **13**:1-11.

doi:10.1186/1752-0509-5-106
Cite this article as: Valgepea *et al.*: Decrease of energy spilling in *Escherichia coli* continuous cultures with rising specific growth rate and carbon wasting. *BMC Systems Biology* 2011 **5**:106.

PUBLICATION III

Peebo K, **Valgepea K**, Nahku R, Riis G, Õun M, Adamberg K & Vilu R

Coordinated activation of PTA-ACS and TCA cycles strongly reduces overflow metabolism of acetate in *Escherichia coli*.

Appl. Microbiol. Biotechnol. 98, 5131–5143, (2014)[1]

[1]Copyright 2014, reprinted with permission from Springer Science and Business Media

Appl Microbiol Biotechnol (2014) 98:5131–5143
DOI 10.1007/s00253-014-5613-y

APPLIED MICROBIAL AND CELL PHYSIOLOGY

Coordinated activation of PTA-ACS and TCA cycles strongly reduces overflow metabolism of acetate in *Escherichia coli*

Karl Peebo · Kaspar Valgepea · Ranno Nahku · Gethe Riis · Mikk Õun · Kaarel Adamberg · Raivo Vilu

Received: 30 October 2013 / Revised: 3 February 2014 / Accepted: 9 February 2014 / Published online: 15 March 2014

Abstract Elimination of acetate overflow in aerobic cultivation of *Escherichia coli* would improve many bioprocesses as acetate accumulation in the growth environment leads to numerous negative effects, e.g. loss of carbon, inhibition of growth, target product synthesis, etc. Despite many years of studies, the mechanism and regulation of acetate overflow are still not completely understood. Therefore, we studied the growth of *E. coli* K-12 BW25113 and several of its mutant strains affecting acetate-related pathways using the continuous culture method accelerostat (A-stat) at various specific glucose consumption rates with the aim of diminishing acetate overflow. Absolute quantitative exo-metabolome and proteome analyses coupled to metabolic flux analysis enabled us to demonstrate that onset of acetate overflow can be postponed and acetate excretion strongly reduced in *E. coli* by coordinated activation of phosphotransacetylase-acetyl-CoA synthetase (PTA-ACS) and tricarboxylic acid (TCA) cycles. Fourfold reduction of acetate excretion (2 vs. 8 % from total carbon) at fastest growth compared to wild type was achieved by deleting the genes responsible for inactivation of acetyl-CoA synthetase protein (*pka*) and TCA cycle regulator *arcA*. The Δ*pka* Δ*arcA* strain did not accumulate any other detrimental by-product besides acetate and showed identical μ_{max} and only ~5 % lower biomass yield compared to wild type. We conclude that a fine-tuned coordination between increasing the recycling capabilities of acetate in the PTA-ACS node through a higher concentration of active acetate scavenging Acs protein and downstream metabolism throughput in the TCA cycle is necessary for diminishing overflow metabolism of acetate in *E. coli* and achieving higher target product production in bioprocesses.

Keywords Acetate overflow · Acs · Continuous culture · Absolute proteomics · Acetylation

Introduction

Acetate is the major by-product in bioprocesses using *Escherichia coli* and glucose as the carbon source. Acetate excretion or overflow during fast growth/glucose uptake aerobic fermentations of *E. coli* has been studied for a long time, and it is known to lead to decreased biomass formation, inhibition of recombinant protein and low molecular weight compound production (Mazumdar et al. 2010; Shiloach et al. 2010; Wolfe 2005).

Acetate overflow is generally believed to be caused by an imbalance between substrate uptake and anabolic/catabolic throughput of downstream pathways, and many theories about different possible limitations in metabolism have been proposed: limitation in tricarboxylic acid (TCA) cycle (Majewski and Domach 1990; Veit et al. 2007), respiratory chain (Han et al. 1992; Paalme et al. 1997; Varma and Palsson 1994) or glyoxylate shunt (GS) activity (Shin et al. 2009; Waegeman et al. 2011) and competition for membrane space (Zhuang et al. 2011). Several process level and genetic manipulations

Electronic supplementary material The online version of this article (doi:10.1007/s00253-014-5613-y) contains supplementary material, which is available to authorized users.

K. Peebo · K. Valgepea · R. Nahku · G. Riis · M. Õun · R. Vilu
Department of Chemistry, Tallinn University of Technology, Tallinn, Estonia

K. Peebo (✉) · K. Valgepea · R. Nahku · G. Riis · M. Õun · K. Adamberg · R. Vilu
Competence Centre of Food and Fermentation Technologies, Akadeemia tee 15A, 12618 Tallinn, Estonia
e-mail: karl@tftak.eu

K. Adamberg
Department of Food Processing, Tallinn University of Technology, Tallinn, Estonia

have been tried to mitigate acetate overflow (reviewed in De Mey et al. 2007). It has been shown that disruption of the main acetate synthesis pathway—phosphotransacetylase-acetate kinase (PTA-ACKA)—reduces acetate excretion but results in lower specific growth rate (μ), biomass yield (Y_{xs}) and substantially elevated lactate and formate excretion (Castaño-Cerezo et al. 2009; Contiero et al. 2000; Dittrich et al. 2005; El-Mansi and Holms 1989; Yang et al. 1999), while deletion of the second major route for acetate production through pyruvate oxidase (PoxB) causes loss of aerobic growth efficiency (Abdel-Hamid et al. 2001). These results demonstrate that acetate overflow cannot be simply avoided by removing the main synthesis pathways without any negative side effects. Furthermore, none of the theories or process/genetic efforts has been able to unequivocally explain the mechanism of acetate overflow.

Recently, a new perspective on the long-studied phenomenon has emerged as it is proposed by several groups that *E. coli* actually synthesises acetate constantly at all μ under aerobic conditions and no acetate overflow occurs at low μ since acetate is fully recycled in the phosphotransacetylase-acetyl-CoA synthetase (PTA-ACS) node (Fig. 1) (Renilla et al. 2011; Shin et al. 2009; Valgepea et al. 2010). Furthermore, we have proposed a new theory for the cause of acetate overflow: acetate excretion in fast growing cells is closely related to the loss of acetate-glucose co-utilisation capability and triggered by carbon catabolite repression-mediated down-regulation of acetyl-CoA synthetase (Acs) leading to disruption of the PTA-ACS node (Valgepea et al. 2010).

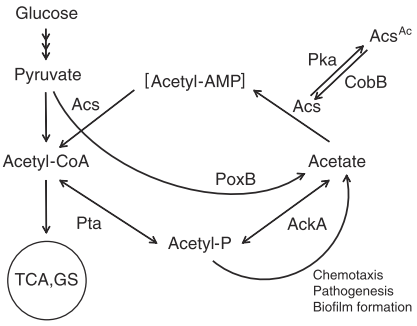

Fig. 1 Simplified overview of glucose metabolism, PTA-ACS node and regulation of Acs by Pka and CobB in *E. coli*. Acetate is mainly generated from: (**a**) acetyl-CoA by Pta and AckA enzymes via the high energy intermediate acetyl-phosphate (acetyl-P); (**b**) acetyl-P during different cellular regulation processes (see Wolfe 2005 and Valgepea et al. 2010 for details); (**c**) pyruvate by PoxB. Acetate in the PTA-ACS node is reactivated by recycling it back to acetyl-CoA by the high-affinity Acs enzyme. Activity of Acs is essential for functioning of the PTA-ACS node: protein lysine acetyltransferase (Pka) inactivates Acs by acetylation (Acs^{Ac}), while the NAD$^+$-dependent protein deacetylase CobB reactivates Acs by deacetylating it. *Pta* phosphotransacetylase, *AckA* acetate kinase, *Acs* acetyl-CoA synthetase, *PoxB* pyruvate oxidase, *TCA* tricarboxylic acid cycle, *GS* glyoxylate shunt

Known regulators involved in the control of *acs* expression in *E. coli* are cAMP receptor protein (CRP), factor for inversion stimulation (Fis), integration host factor (IHF) (Wolfe 2005) and possibly the dual transcriptional regulator for anoxic redox control (ArcA) (Covert et al. 2004). Besides transcriptional regulation, Acs protein activity in bacteria is regulated by the Pat/CobB system which was first described in *Salmonella enterica* (Starai et al. 2002; Starai and Escalante-Semerena 2004). An identical system is present in *E. coli*: protein lysine acetyltransferase (Pka) inactivates Acs by acetylation (Castaño-Cerezo et al. 2011), while the NAD$^+$-dependent regulator protein deacetylase CobB releases Acs from repression by deacetylating it (Zhao et al. 2004) (Fig. 1). Interestingly, increased Acs activity together with derepression of GS enzymes decreases acetate accumulation in *E. coli* batch cultures (Shin et al. 2009) pointing to the importance of Acs and its active/inactive pools in acetate metabolism.

As regulation of Acs and PTA-ACS node seems to have an important role in acetate overflow, the aim of this work was to gain further understanding of the relevance of the PTA-ACS node in overflow metabolism possibly leading to reduction of acetate overflow in aerobic *E. coli* cultivations. For this, we studied the growth of *E. coli* K-12 BW25113 and several of its mutant strains affecting acetate-related pathways in continuous cultures (accelerostat (A-stat); Paalme et al. 1995) at various specific glucose consumption rates (q_{glc}) using absolute quantitative exo-metabolome and proteome analyses coupled to metabolic flux analysis (MFA). Furthermore, as repression of the high-affinity acetate scavenging enzyme Acs seems to be central in acetate overflow regulation through the loss of acetate-glucose co-utilisation capability (Valgepea et al. 2010), we focused this study on Acs-related inhibition of acetate overflow by improving acetate recycling capabilities.

We chose A-stat (Paalme et al. 1995) for the cultivation method since it allows accurate detection of metabolic switch points (e.g. start of acetate overflow) and high-resolution description of metabolism dynamics (e.g. acetate accumulation) which are both instrumental for a system-wide study of acetate overflow regulation in *E. coli* (Valgepea et al. 2010). A-stat experiments produce comparable results with chemostats at the level of major growth characteristics in *E. coli* and *Lactococcus lactis* (Lahtvee et al. 2011; Valgepea et al. 2010, 2011) and at transcriptome (Valgepea et al. 2010) and proteome (Nahku 2012) levels in *E. coli*.

In this work, we show that acetate overflow in aerobic *E. coli* cultivations can be remarkably reduced by coordinated activation of PTA-ACS and TCA cycles which was achieved by deleting the Acs-inactivating gene *pka* and TCA cycle regulator *arcA*. We conclude that a fine-tuned coordination between increasing the recycling capabilities of acetate in the PTA-ACS node through a higher concentration of active acetate scavenging Acs protein and downstream metabolism

throughput in the TCA cycle is necessary for diminishing overflow metabolism of acetate in *E. coli*.

Materials and methods

Bacterial strains, growth medium and continuous culture conditions

Bacterial strains used in this work were *E. coli* K-12 BW25113 wild type (WT) and its single-gene knockouts obtained from the Keio collection (Baba et al. 2006) or constructed using homologous recombination. Strains overexpressing *acs* (*acs*OE and Δ*pka acs*OE) were constructed by swapping the native *acs-yjcH-actP* promoter region of the parent strain (WT and Δ*pka*, respectively) with a P*tet* promoter sequence (Veit et al. 2007). The P*tet* promoter sequence was first annealed with a kanamycin resistance gene from a pCP13 plasmid (Datsenko and Wanner 2000) by PCR after which this construct was used to swap the native promoter region using homologous recombination as described by Datsenko and Wanner (2000). For the construction of Δ*pka* Δ*arcA* double knockout strain, *arcA* gene was deleted from the Δ*pka* strain using a homologous recombination method that uses pSIM6 plasmid instead of pKD46 (Sharan et al. 2009). The strains constructed in this work are deposited in the public strain collection of the Collection of Environmental and Laboratory Strains (University of Tartu, Estonia) with reference numbers shown in Table 1. All strains used in this work with their abbreviations are listed in Table 1.

All A-stat experiments were carried out using a defined minimal medium supplemented with glucose (4.5 g/L) as described before by Valgepea et al. (2010). Glucose-limited cultures were grown in 1.25 L BioBundle bioreactors (Applikon Biotechnology B.V., Schiedam, the Netherlands) controlled by an ez-Control biocontroller (Applikon Biotechnology B.V.) and a cultivation control program "BioXpert XP" (Applikon Biotechnology B.V.). The system was equipped with OD, pH, pO_2, off-gas O_2 and CO_2 and temperature sensors.

A-stat experiments were carried out at 37 °C, pH 7 and aerobic conditions ensured by agitation of 800 rpm and air flow rate of 150 mL/min. Glucose-limited cultures were stabilised in chemostat at dilution rate (D) 0.1 or 0.2 h^{-1} after which A-stat with a continuous increase of D (acceleration of 0.01 h^{-2}) was initiated. The control algorithm for A-stat was: $D=D_0+a_D \times t$, where D_0 is the initial D of chemostat (h^{-1}), a_D is the acceleration of the A-stat phase (h^{-2}) and t is the time from the start of A-stat (h).

WT and Δ*arcA*, Δ*pka* Δ*arcA* strains were cultivated in four and two independent A-stat experiments, respectively. Other strains were cultivated in single A-stat experiments. Growth characteristics in A-stat were calculated using the formulas described in a previous study (Kasemets et al. 2003) on the basis of total volume of medium pumped out from bioreactor (L), biomass (gram dry cellular weight (gDCW)/L) and organic acid concentrations (mM) in culture broth and CO_2 concentration in the off-gas (mM). We note that in CO_2 production calculations, also the CO_2 and HCO_3^- dissolved in the culture broth were taken into account as described by Taymaz-Nikerel et al. (2013).

Analytical methods

Exo-metabolome analysis for a more complete carbon balance was carried out as described previously in Valgepea et al. (2011). In short, samples of culture broth were centrifuged (14,000×*g* for 5 min), and the supernatant was collected and analysed for glucose and organic acids by HPLC (Alliance2795 system, Waters Corporation, Milford, MA, USA) using either a Bio-Rad HPX-87H Aminex ion-exclusion (Bio-Rad Laboratories, Inc., Hercules, CA, USA) or Agilent Hi-Plex H (Agilent Technologies, Santa Clara, CA, USA) column connected to RI and UV detectors (35 °C, flow rate 0.6 mL/min). The column was eluted with 4.1 mM sulphuric acid for glucose, carbamoyl-aspartate, lactate and orotate and with

Table 1 *E. coli* K-12 strains used in this work

Strain	Genotype	Reference	Abbreviations
BW25113	*lacI*^{q} *rrnB3* Δ*lacZ*4787 *hsdR*514 Δ(*araBAD*)567 Δ(*rhaBAD*)568 *rph-1*	Datsenko and Wanner (2000)	WT
JW4030	[BW25113] Δ*acs*::*kan*	Baba et al. (2006)	Δ*acs*
JW2568	[BW25113] Δ*pka*::*kan*	Baba et al. (2006)	Δ*pka*
JW1106	[BW25113] Δ*cobB*::*kan*	Baba et al. (2006)	Δ*cobB*
JW4364	[BW25113] Δ*arcA*::*kan*	Baba et al. (2006)	Δ*arcA*
RV02	[BW25113] *acsp1*, *acsp2*::P*tet*	This work	*acs*OE
RV03	[BW25113] Δ*pka*::*kan*; *acsp1*, *acsp2*::P*tet*	This work	Δ*pka acs*OE
RV04	[BW25113] Δ*pka*::*kan*; Δ*arcA*::*kan*	This work	Δ*pka* Δ*arcA*

26.5 mM formic acid for acetate, dihydroorotate and acetyl-aspartate analysis.

Biomass concentration (X) in the experiments with WT, Δ*arcA* and Δ*pka* Δ*arcA* was determined gravimetrically and expressed in gram dry cell weight per liter. Approximately 15 ml of culture broth was collected, weighed and transferred onto a pre-weighed glass fiber filter (S-Pak Membrane Filter White gridded 0.45 μm, EMD Millipore Corporation, Billerica, MA, USA) which was mounted on a vacuum filtration device. Special care was taken to ensure that all the culture broth was transferred onto the filter. Next, the filter was heated in an oven at 105 °C for >24 h after which the filter with dried biomass was cooled in a desiccator for >1 h and finally weighed to determine the dry biomass weight. Coefficient of variation for the determination of X with this method between technical replicates was <1 % (three technical replicates in 27 sample points). The developed method was used to determine the correlation factor (K) for all μ values between X and optical density (at 600 nm) of culture broth (OD), ($K=X/\mathrm{OD}$). The obtained linear μ-dependent relationship of $K=-0.2309\times\mu+0.4534$ for WT was used to calculate X for the experiments with strains Δ*cobB*, Δ*acs*, Δ*pka*, *acs*OE and Δ*pka acs*OE based on measured OD values.

Metabolic flux analysis

MFA using strain-dependent growth characteristics and exometabolome data for Δ*arcA*, Δ*pka* Δ*arcA* and WT was carried out as described previously (Valgepea et al. 2011) except that pyruvate instead of oxaloacetate was set as the metabolite characterising carbon outflow that was not identified experimentally. In short, the simplified metabolic network contained the main central carbon metabolism pathways—glycolysis, TCA cycle, pentose phosphate pathway (PPP)—one pyrimidine synthesis pathway (to include the excreted carbamoyl-aspartate, orotate and dihydroorotate) and the acetyl-aspartate synthesis reaction. The reconstructed network converted into a fully determined and calculable stoichiometric matrix consisting of 24 metabolites and 50 fluxes (24 unknown, 1 measured inflow, 7 outflow and 18 calculated fluxes based on biomass composition and stoichiometries of anabolic pathways). The same metabolic network, stoichiometric matrix and other MFA parameters were used for all the three analysed strains, since, to the best of our knowledge, neither *arcA* nor *pka* deletion should affect MFA characteristics compared to WT.

Proteome analysis

Sample preparation

One milliliter of WT, Δ*pka*, Δ*arcA* and Δ*pka* Δ*arcA* culture broth was harvested at $\mu=0.4\ \mathrm{h}^{-1}$ and centrifuged (14,000×*g* for 1 min at 4 °C), and the pellet was washed once with phosphate buffered saline (PBS), flash frozen with liquid N_2 and kept at −80 °C until further processing. Frozen pellets were melted on ice after which 100 μg of the sample biomass was mixed with 100 μg of stable isotope labeling by amino acids in cell culture (SILAC)-labeled (Ong et al. 2002) *E. coli* biomass (internal standard) and frozen at −80 °C. Cell pellets were suspended in 100 μL SDS lysis buffer (4 % SDS/100 mM Tris-HCl pH 8/100 mM DTT) and heated at 95 °C for 5 min. Cell lysates were sonicated with ultrasound for a few pulses and pelleted by centrifugation. Cell lysates were digested with trypsin according to the Filter Aided Sample Preparation protocol (FASP) (Wiśniewski et al. 2009) and purified with C-18 StageTips (Rappsilber et al. 2007).

Sample analysis

Sample analysis was done as in Arike et al. (2012) with the exception of running one technical replicate for WT cells (triplicate for mutant strains). In short, LC-MS/MS analysis was performed using an Agilent 1200 series nanoflow system (Agilent Technologies) connected to an LTQ Orbitrap mass spectrometer (Thermo Electron, San Jose, CA, USA) equipped with a nanoelectrospray ion source (Proxeon, Odense, Denmark). Peptides were separated with a 240-min gradient from 2 to 40 % B (A—0.5 % acetic acid, B—0.5 % acetic acid/80 % acetonitrile) using a flow rate of 200 nL/min. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaíno et al. 2013) with the dataset identifier PXD000556.

Data analysis

Data analysis of raw MS files was performed by the MaxQuant software package version 1.3.0.5 (Cox and Mann 2008). Peak lists were searched using the Andromeda search engine (built into MaxQuant) against an *E. coli* database (downloaded on September 5, 2013 from http://www.uniprot.org/) which was supplemented with common contaminants (e.g. human keratin, trypsin). Full tryptic specificity, a maximum of two missed cleavages and a mass tolerance of 0.5 Da for fragment ions, was specified in the MaxQuant search. Carbamidomethylation of cysteine was set as a fixed modification, and methionine oxidation and protein N-terminal acetylation were set as variable modifications. The required false discovery rate was set to 1 % for both peptide and protein levels and the minimum required peptide length was set to seven amino acids. “Match between runs” option with a time window of 2 min was allowed. To increase peptide identification by “Match between runs” function, additional 26 WT *E. coli* raw proteome MS files were supplemented to the MaxQuant search.

Relative and absolute proteome quantification

Since each sample was mixed with equal amounts of SILAC-labeled *E. coli* internal standard, relative protein expression differences for >1,000 proteins between mutants and WT could be calculated from expression differences between each sample and the internal standard. All protein ratios were normalised assuming a median expression change of 1 since equal amounts of labeled internal standard and unlabeled sample were mixed together.

Absolute proteome quantification to determine intracellular protein concentrations (protein copies/fL of biomass) was performed using a label-free protein quantification method termed the total protein approach (Arike et al. 2012; Wiśniewski et al. 2012). This method assumes that most of the proteins are detected by MS and this is usually the case as the top 1,000 proteins make up more than 90 % of the protein mass (Milo 2013; Valgepea et al. 2013). For each protein, its relative fraction from the total proteome can be derived from the ratio of its individual LFQ intensity (reported by MaxQuant) to the total MS signal. This ratio can be converted into protein concentration (copies/fL) by dividing with the protein's molecular weight and multiplying with the Avogadro constant and the protein concentration in a cell (g protein/mL cell volume) (Milo 2013). We determined the protein concentration using measured biomass concentration and protein fraction of dry biomass, cell buoyant density of 1.1 g/mL (Baldwin et al. 1995; Loferer-Krössbacher et al. 1998) and dry fraction of wet biomass of 30 % (Cayley et al. 1991). As SILAC-labeling is more sensitive and reproducible than label-free quantification (Asara et al. 2008; Hendrickson et al. 2006), we determined protein concentrations using the above-described method only for a WT sample from the A-stat experiment at μ=0.4 h^{-1}. All the reported protein concentrations in this work were calculated by multiplying the values of the latter sample with the relative protein expression differences between each sample and the SILAC-labeled *E. coli* internal standard to maintain the high accuracy of SILAC labeling for relative comparison of strains. Protein concentrations instead of abundances per cell were calculated because cell size can vary between conditions and strains (Milo 2013; Valgepea et al. 2013).

Gene expression analysis

Real-time PCR for *acs* expression analysis was carried out as described before (Nahku et al. 2010) except that the primer concentration used in this work was 20 μM.

Results

Glucose-limited cultures were first stabilised in a chemostat at either D=0.1 or 0.2 h^{-1} to achieve steady-state conditions after which the A-stat phase with a continuous increase of D was started using an acceleration of 0.01 h^{-2}. This yielded steady-state representative growth until cells could not keep up with the rising D after achieving μ_{max}, resulting in culture washout. *E. coli* BW25113 wild-type (WT) and Δ*arcA*, Δ*pka* Δ*arcA* strains were cultivated in four and two independent A-stat experiments, respectively. Other strains were analysed in single experiments since the data of strains cultivated in multiple replicates confirmed the high reproducibility of A-stat data (Lahtvee et al. 2011; Valgepea et al. 2010). We note that growth characteristics of different strains are compared relative to q_{glc} instead of μ in this study since glucose uptake is known to regulate the magnitude of carbon catabolite repression (Görke and Stülke 2008; Narang 2009) possibly triggering the down-regulation of Acs and acetate overflow (Valgepea et al. 2010).

Effect of Acs deletion and overexpression on acetate metabolism

We first wanted to verify the hypothesis that *E. coli* synthesises acetate also at low μ using WT with deleted *acs* (Δ*acs*) which should lead to disruption of the PTA-ACS node and loss of acetate recycling. Indeed, as expected, the PTA-ACS node disruption in Δ*acs* led to acetate excretion at all q_{glc} values studied, while no acetate accumulated in WT up to q_{glc}~4.2 mmol $gDCW^{-1}$ h^{-1} (Fig. 2a). This phenotype of Δ*acs* is consistent with previous observations (Renilla et al. 2011; Shin et al. 2009) and confirms the hypothesis that acetate is constantly synthesised at all μ (Renilla et al. 2011; Shin et al. 2009; Valgepea et al. 2010). As Δ*acs* constantly directed more carbon to acetate, a lower maximum q_{glc} (Fig. 2a) and on average 5 % lower Y_{xs} (0.42±0.00 vs. 0.44 ±0.01 gDCW/g glucose; average of the q_{glc} range±standard deviation) were observed compared to WT.

As one of the proposed reasons for acetate overflow is the repression of Acs (Valgepea et al. 2010), we tested the effect of increasing Acs expression on acetate excretion. Hence, *acs* was overexpressed (*acs*OE) by swapping the native promoter of the *acs-yjcH-actP* operon by a strong constitutive promoter P*tet* (Veit et al. 2007) to maintain recycling of acetate in the PTA-ACS node also at faster growth. A 114±10.3- and 9±0.6-fold higher expression of *acs* messenger RNA (mRNA) in *acs*OE compared to WT was measured by real-time PCR at μ=0.5 h^{-1} (q_{glc}~6 mmol $gDCW^{-1}$ h^{-1}) and μ=0.2 h^{-1} (q_{glc}~3 mmol $gDCW^{-1}$ h^{-1}), respectively. However, acetate overflow started at a lower q_{glc} in *acs*OE, and the strain also excreted more carbon as acetate than WT at all the q_{glc} studied (Fig. 2a). Similar to Δ*acs*, lower maximum q_{glc} and μ_{max} were achieved in *acs*OE compared to WT. The fact that higher expression of the acetate-consuming gene *acs* led to earlier onset of acetate overflow and higher carbon wasting to acetate could possibly be a result of either the negative effects of too

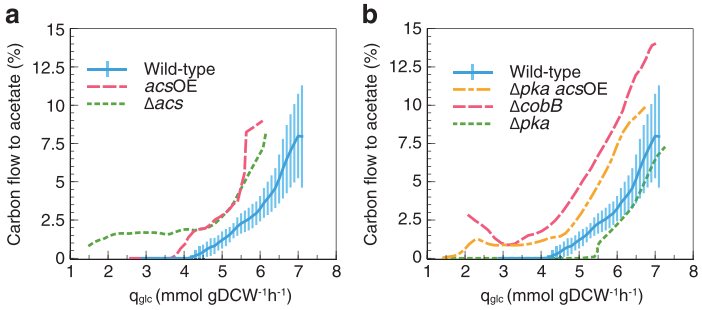

Fig. 2 Specific glucose consumption rate (q_{glc})-dependent carbon flow to acetate in *E. coli* K-12 BW25113 wild type and its Acs-related mutants. The percentage of carbon flow from glucose to acetate: **a** in wild type (*blue solid line*), *acs*OE (*pink dashed line*), Δ*acs* (*green dotted line*); **b** in WT (*blue solid line*), Δ*cobB* (*pink dashed line*), Δ*pka* (*green dotted line*), Δ*pka acs*OE (*yellow dash-dotted line*). *Lines* for each strain are best-fit splines of 9–36 acetate and glucose concentration measurements in each experiment. *Error bars* represent standard deviation of four independent A-stat experiments

high protein overexpression (see "Discussion") and/or due to most of the Acs protein pool being inactive. Hence, we next concentrated on the effects regulating Acs protein activity.

Protein acetylation plays an important role in acetate metabolism

Two enzymes in *E. coli* are known to regulate the activity of Acs protein: Pka inactivates Acs by acetylation (Castaño-Cerezo et al. 2011), while CobB can release Acs from repression by deacetylating it (Zhao et al. 2004) (Fig. 1). Moreover, it has been shown in vitro that Acs can autoacetylate itself in the presence of acetate and ATP (Barak et al. 2004), further demonstrating the importance of acetylation in determining Acs protein activity. Hence, to investigate the effects of altered protein acetylation on acetate metabolism, we analysed growth of Pka (Δ*pka*) and CobB (Δ*cobB*) single-gene knockout mutants of *E. coli*.

Δ*cobB* cells, which cannot reactivate Acs by deacetylation, excreted more acetate (Fig. 2b) and showed on average 28 % lower Y_{xs} (0.32±0.01 vs. 0.44±0.01 gDCW/g glucose) compared to WT at all studied q_{glc} values. Similar to the Δ*acs* strain, a relatively constant fraction from the carbon balance was directed to acetate up to q_{glc}~4 mmol gDCW^{-1} h^{-1} (Fig. 2b). This further confirms the importance of the state of the Acs protein pool and an active PTA-ACS node in acetate metabolism.

Next, we analysed the growth of a strain lacking Pka (Δ*pka*), known to inactivate Acs in *S. enterica* (Starai and Escalante-Semerena 2004) and *E. coli* (Castaño-Cerezo et al. 2011), which should eliminate or at least substantially decrease the fraction of inactivated Acs protein in the cell, possibly leading to higher recycling of acetate in the PTA-ACS node and postponed acetate overflow. Indeed, Δ*pka* showed a postponed start of acetate overflow at q_{glc}=5.5 compared to ~4.2 mmol gDCW^{-1} h^{-1} in WT (Fig. 2b). We wanted to verify if the diminished acetate overflow in Δ*pka* was indeed caused by higher Acs activity in vivo. However, this hypothesis cannot be unfortunately tested since, to the best of our knowledge, there exists no assay to unequivocally determine the in vivo activity of Acs in *E. coli* cells expressing proteins of both acetate-consuming pathways (ACKA-PTA and Acs). Since all the strains studied in this work express the three enzymes of the acetate-consuming pathways (AckA, Pta and Acs) at considerable concentrations as can be seen from the absolute quantitative proteome data described below, in vivo activity of Acs cannot be unequivocally determined in any of the strains.

As we saw a positive effect of *pka* deletion on diminishing acetate metabolism, we next introduced *pka* deletion into the *acs*OE background (Δ*pka acs*OE) to test if the lack of a positive effect of *acs* overexpression on mitigating acetate accumulation could have arisen from a substantial part of the Acs protein pool being inactive in the *acs*OE strain (see above). This was not the case as acetate overflow was not postponed in Δ*pka acs*OE compared to WT (Fig. 2b). Therefore, we searched for other means to further postpone the onset and decrease the magnitude of acetate overflow.

Strong reduction of acetate overflow by higher levels of active Acs and TCA cycle throughput

Our results showed that strong overexpression of *acs* had no positive effects on relieving acetate overflow. However, since we saw diminished overflow of acetate in Δ*pka* which altered the levels of active Acs, we surmised that acetate overflow could be further reduced in Δ*pka* if Acs levels would be slightly higher, but not too high as in Δ*pka acs*OE (see "Discussion"). Furthermore, it could be plausible that even if the Δ*pka* cells could recycle more acetate due to higher levels of active Acs, downstream pathways (e.g. TCA cycle) could still be limiting in processing the recycled acetyl-CoA to

divert carbon away from acetate. Hence, we also looked for ways to increase downstream throughput for the recycled acetate.

Deleting ArcA induces expression of TCA cycle genes (Covert et al. 2004; Vemuri et al. 2006) and flux throughput (Perrenoud and Sauer 2005). Interestingly, Covert et al. (2004) also detected ~2-fold higher expression of *acs* in the *arcA* deletion mutant compared to WT. Hence, it seemed that an *arcA* deletion would hit two birds with one stone for us, and therefore, we deleted *arcA* in the Δ*pka* background (Δ*pka* Δ*arcA*) in order to concurrently increase the levels of active Acs and downstream throughput of the TCA cycle to further diminish acetate overflow. Remarkably, the double-knockout strain showed even further reduced acetate overflow as the onset of acetate accumulation was postponed until q_{glc}~ 6 mmol gDCW^{-1} h^{-1} (q_{glc}~4.2 mmol gDCW^{-1} h^{-1} in WT), and more importantly, carbon wasting into acetate in Δ*pka* Δ*arcA* was more than 4-fold lower compared to WT (2 vs. 8 % of total carbon) at maximal q_{glc} (Fig. 3a).

Carbon balance showed that Δ*pka* Δ*arcA* primarily produced more CO_2 instead of acetate (2–8 % more than WT) (Fig. 3b). In addition, MFA revealed that the higher CO_2 production was the result of increased TCA cycle fluxes as expected from an *arcA* deletion (Fig. 3c; Supplementary Table S1). MFA also showed that while more CO_2 was produced in the TCA cycle, less CO_2 was produced through PPP in Δ*pka* Δ*arcA*, demonstrating the relevance of balancing TCA cycle and PPP also in acetate metabolism (Fig. 3d; Supplementary Table S1). Compared to acetate overflow-reduced strains of acetate synthesis pathway deletions (*pta*+ *ackA* and *poxB*) which show reduced μ, Y_{xs} and substantially elevated lactate and formate excretion (see above), our Δ*pka* Δ*arcA* strain with strongly reduced overflow did not accumulate any other detrimental by-product, maintained μ_{max} of WT (Fig. 3a) and exhibited only ~5 % lower Y_{xs} (0.42±0.00 vs. 0.44±0.01 gDCW/g glucose) compared to WT.

The results of the study of Vemuri et al. (2006) indicated that acetate overflow could be postponed by a single *arcA* knockout in *E. coli* K-12 MG1655. Hence, we also analysed the growth of Δ*arcA* in our *E. coli* K-12 BW25113 background to check if the *arcA* knockout effect could be strain-dependent and whether the positive effects of Δ*pka* Δ*arcA* on diminishing acetate overflow could actually arise only from an *arcA* deletion. It is evident that the *arcA* deletion alone could not be accounted for the reduced overflow in Δ*pka* Δ*arcA* in our BW25113 background as acetate overflow started even

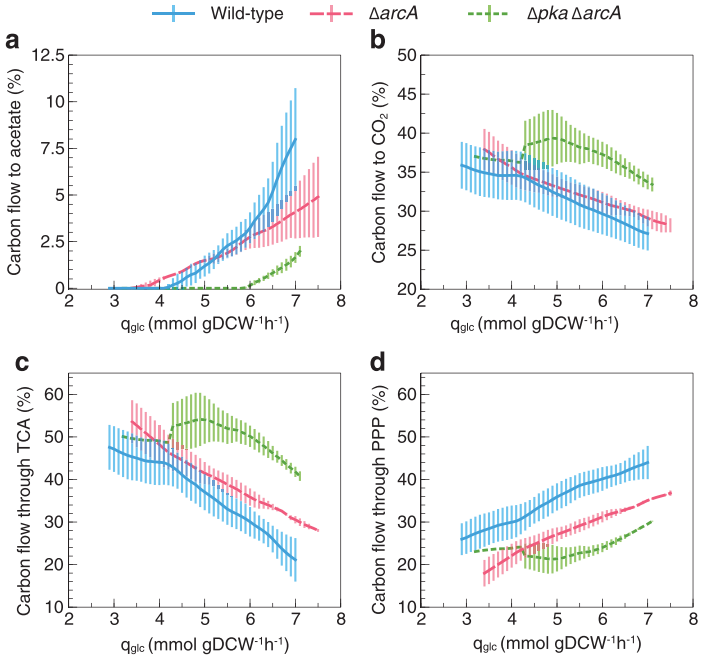

Fig. 3 Specific glucose consumption rate (q_{glc})-dependent carbon flow through metabolism in *E. coli* K-12 BW25113 wild type and its acetate overflow-reduced mutants. The percentage of carbon flow from glucose to acetate (**a**) and CO_2 (**b**). The percentage of carbon flow from glucose through TCA cycle (**c**) and pentose phosphate pathway (PPP) (**d**) represented by suc and gnd fluxes, respectively, expressed as the C-molar percentage of flux through the reaction from the consumed carbon. Wild type (*blue solid line*), Δ*arcA* (*pink dashed line*), Δ*pka* Δ*arcA* (*green dotted line*). *Lines* for each strain are best-fit splines of 13–43 glucose and product concentration measurements in each experiment. *Error bars* represent standard deviation of four and two independent A-stat experiments for wild type and Δ*arcA*, Δ*pka* Δ*arcA*, respectively

earlier (q_{glc}~3.5 vs. ~4.2 mmol gDCW^{-1} h^{-1}) in Δ*arcA* compared to WT (Fig. 3a). Interestingly, accumulation of acetate was slower and Δ*arcA* wasted ~4 % less carbon into acetate compared to WT near maximal q_{glc} (Fig. 3a). As expected, Δ*arcA* diverted less carbon to CO_2 compared to Δ*pka* Δ*arcA* (Fig. 3b; Supplementary Table S1), which is possibly due to an active PTA-ACS node in Δ*pka* Δ*arcA* that provides the TCA cycle with more acetyl-CoA through higher acetate recycling compared to Δ*arcA*. The latter points to the importance of Δ*pka* in facilitating higher TCA cycle throughput by supplying it with more acetyl-CoA through higher acetate recycling in the PTA-ACS node.

Proteomic response to different genetic disruptions

We carried out absolute quantitative proteomics analysis of WT and several mutant strains yielding intracellular protein concentrations to find out if the different metabolic flux patterns could be described by proteome expression profiles. For this, we used SILAC-labeled *E. coli* biomass as an internal standard (see "Materials and methods") to compare the proteomes in single A-stat experiments of Δ*pka*, Δ*arcA* and Δ*pka* Δ*arcA* with WT at μ=0.4 h^{-1} (Fig. 4 and Supplementary Table S2). We chose to compare the proteome expression profiles at μ=0.4 h^{-1} corresponding to q_{glc} values of 5.55, 5.25, 5.44 and 5.01 mmol gDCW^{-1} h^{-1} for Δ*pka*, Δ*arcA*, Δ*pka* Δ*arcA* and WT, respectively, since significant differences in acetate overflow characteristics between the strains had occurred by that high μ (q_{glc}). Proteome analysis for WT was carried out also at μ=0.2 and 0.51 h^{-1} in two independent A-stat experiments yielding an average coefficient of variation of protein expression changes over the quantified >1,000 proteins to be <7 %. The latter means that protein expression differences >14 % between a mutant strain and WT could be considered significant. Proteome-wide data for protein concentrations and expression changes between strains is in Supplementary Table S2.

Based on the observation that *acs* expression is increased ~2-fold in Δ*arcA* compared to WT (Covert et al. 2004), we introduced *arcA* deletion into Δ*pka* strain with the expectation of increasing the levels of active Acs protein. Indeed, our proteome analysis confirmed the latter as expression of Acs increased ~1.7-fold in response to *arcA* deletion both in Δ*arcA* and Δ*pka* Δ*arcA* strains (Fig. 4; all comparisons relative to WT). This demonstrates that the total Acs protein concentration, most probably together with the levels of its active form, plays a role in acetate metabolism as acetate overflow was strongly reduced in Δ*pka* Δ*arcA* (see above).

We detected increased carbon fluxes through the TCA cycle resulting in higher CO_2 fraction from the carbon balance in Δ*pka* Δ*arcA* (Fig. 3b, c). This was confirmed also on the translational level as expression of the TCA cycle enzymes was enhanced in Δ*pka* Δ*arcA* (on average 1.19-fold compared to WT; p~0.001, paired two-tailed t test) (Fig. 4). Notably, although TCA cycle fluxes in Δ*arcA* were lower compared to Δ*pka* Δ*arcA* at the proteome data points of q_{glc}~5 mmol gDCW^{-1} h^{-1}, protein concentrations in the pathway were elevated in Δ*arcA* (on average 1.25-fold, $p<10^{-3}$). The result of lower flux throughput with higher protein concentrations (comparing Δ*arcA* to Δ*pka* Δ*arcA*) points to non-constant catalytic rates of enzymes exerted through post-translational regulation. This observation is in line with recent bacterial studies (Adamberg et al. 2012; Valgepea et al. 2013) which demonstrate that post-translational control of fluxes through alteration of catalytic rates of enzymes is the prevalent mode of flux regulation in *E. coli* and *L. lactis*.

Deletion of *arcA* leads to a strong 7-fold expression increase of the TCA cycle and respiratory chain mRNAs in *E. coli* chemostat cultures at D=0.4 h^{-1} (Vemuri et al. 2006). However, protein concentrations in the TCA cycle were increased only 1.50-fold in our Δ*arcA* strain at the same D (Fig. 4). Moreover, instead of a 7-fold increase, we detected lower expression of cytochromes (CyoABD and CydA) both in Δ*arcA* and Δ*pka* Δ*arcA* and small changes in NADH dehydrogenase complexes NDH I and NDH II (Fig. 4). The difference between our results and those of Vemuri et al. (2006) highlights the importance of protein measurements as analysing only mRNA data could lead to equivocal conclusions.

Shin et al. (2009) report that expression of GS in addition to Acs has to be increased for reducing acetate overflow in *E. coli* batch cultures. However, we detected lower expression of GS proteins AceA and AceB in Δ*pka* Δ*arcA* (Fig. 4), and yet this strain showed strongly reduced acetate overflow. This discrepancy could possibly be explained by different *E. coli* strains (W3110 in Shin et al. 2009) and growth conditions of glucose limitation in our case and glucose excess in the experiments by Shin et al. (2009).

Regarding other possibly relevant protein expression changes for reduced acetate overflow, expression of PoxB (converts pyruvate and ubiquinone to acetate and ubiquinol) was considerably lower in all analysed mutants (Fig. 4). Although this could be relevant for lower acetate accumulation at faster growth since PoxB is suggested to be the main enzyme synthesising acetate at high μ (Nahku et al. 2010), the protein concentration of PoxB in all the mutants was low compared to other PTA-ACS node proteins (Supplementary Table S2). In addition, we detected a very strong up-regulation, 2–6-fold in Δ*pka* Δ*arcA* and 5–30-fold in Δ*pka* and Δ*arcA*, of proteins related to chemotaxis and flagellar assembly (Supplementary Table S2). However, this has been seen before: highly frequent insertions of insertion sequence elements to the regulatory region of *fliD* cause expression changes of these genes (Barker et al. 2004; Ishii et al. 2007; Nahku et al. 2011).

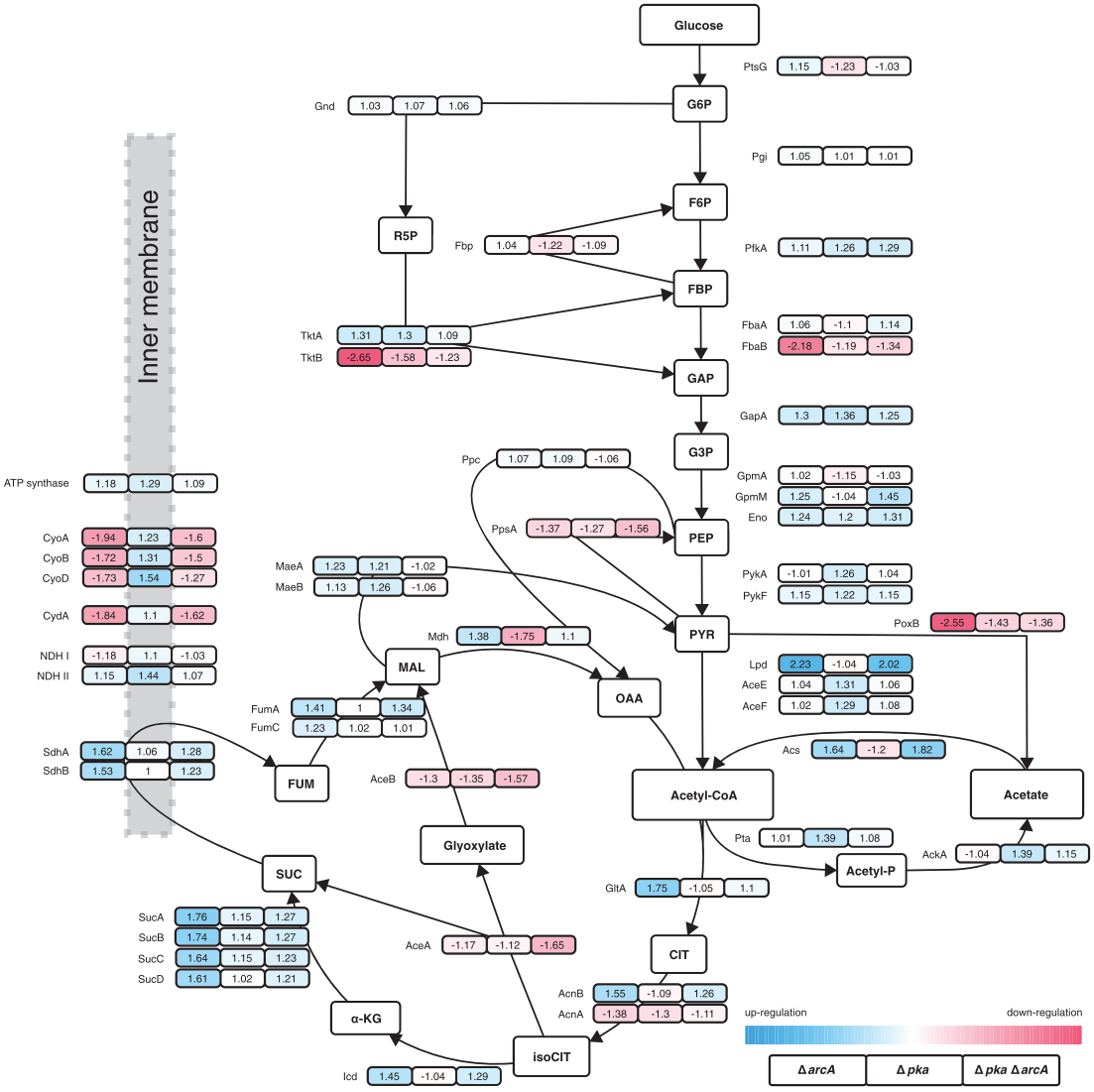

Fig. 4 Proteome expression of *E. coli* K-12 BW25113 Δ*arcA* (*left box*), Δ*pka* (*middle box*) and Δ*pka* Δ*arcA* (*right box*) compared to wild type at μ=0.4 h^{-1}. *Numbers* represent fold change of protein concentration compared to wild type; down-regulation is shown in *red* and denoted by negative values, up-regulation in *blue* and by positive values. Colour intensity is proportional to fold change and scaled by the maximum and minimum protein concentration fold change on the figure. Proteins aligned on the *grey area* are known to be associated with cell membrane. NDH I and ATP synthase represent the average fold change of NADH dehydrogenase and ATP synthase complex protein concentrations, respectively. *Arrow heads* show the presumed reaction direction in given conditions

Discussion

In this work, we studied the growth of *E. coli* K-12 BW25113 and several of its mutant strains affecting acetate-related pathways with the aim of further understanding the relevance of the PTA-ACS node in overflow metabolism and mitigating acetate overflow in aerobic *E. coli* cultivations. Firstly, our results confirm the notion that acetate is synthesised in vivo also at low μ (Renilla et al. 2011; Shin et al. 2009; Valgepea et al. 2010) as strains defective in acetate re-utilisation (Δ*acs* and Δ*cobB*) excreted acetate during slow growth where WT did not (Fig. 2a, b). This further supports the idea that an

active Acs pool prevents acetate overflow during slow growth of WT *E. coli* since acetate is fully recycled in the PTA-ACS node. Moreover, the PTA-ACS node is also relevant for growth efficiency since the higher acetate excretion in both of the mutants with disrupted PTA-ACS node (Δ*acs* and Δ*cobB*) leads to lower Y_{xs} compared to WT. Both Δ*acs* and Δ*cobB* mutations seem to have further negative effects on Y_{xs} in addition to carbon loss to acetate as the increase of carbon flux to acetate cannot solely explain the lower Y_{xs} values observed. For example, the substantially larger fraction of acetate in the carbon balance of Δ*cobB* compared to WT at high q_{glc} values (Fig. 2b) shows that Acs is not the only protein specifically regulated by CobB, as also shown before (Weinert et al. 2013).

An intermediate molecule of the PTA-ACS cycle—acetyl-phosphate (acetyl-P)—plays a critical role in the acetylation of many proteins (Weinert et al. 2013) and regulation of chemotaxis (Barak et al. 1998; Klein et al. 2007; Mayover et al. 1999), pathogenesis (Anfora et al. 2008), protein degradation (Mizrahi et al. 2009), etc. in *E. coli* (Fig. 1). Based on the experiment with Δ*acs*, we propose that ~0.8 mmol $gDCW^{-1}$ h^{-1} (corresponding to 2 % from the carbon balance in Δ*acs*) of acetate is constantly recycled in the PTA-ACS node of WT *E. coli* to enable rapid regulation of the acetyl-P pool. This value is higher compared to the calculated value of ~0.3 mmol $gDCW^{-1}$ h^{-1} at μ=0.2 h^{-1} based on chemotaxis needs and in vivo synthesis of acetate as a by-product during synthesis of cysteine, methionine and arginine (Valgepea et al. 2010), possibly due to the fact that requirement of acetyl-P for protein acetylation was not taken into account in these calculations.

The fact that lack of (active) Acs causes acetate overflow at low μ (Fig. 2a) and the hypothesis that acetate accumulation with increasing q_{glc} is triggered by Acs repression (Valgepea et al. 2010) obviously lead to the idea that higher expression of Acs should reduce acetate overflow. Although *acs* overexpressing strains have been constructed (Lin et al. 2006; Shin et al. 2009), no steady-state data at different q_{glc} values exists. Moreover, the data in the literature is somewhat contradictory. Shin et al. (2009) showed that both higher activity of GS and *acs* overexpression are necessary, whereas only higher expression of Acs was concluded to be sufficient by San and coworkers (Lin et al. 2006) to decrease acetate accumulation in *E. coli* batch cultures. We, however, saw no reduction of acetate overflow in the constructed *acs*OE strain in our continuous cultures (Fig. 2a). Furthermore, the possibility of inactivation of the higher Acs pool in *acs*OE was ruled out by introducing deletion of the gene—*pka*—responsible for Acs inactivation in *E. coli* (Castaño-Cerezo et al. 2011) in the latter strain since the double mutant strain (Δ*pka acs*OE) accumulated more acetate than WT (Fig. 2b). As both *acs* overexpression mutants accumulated more acetate than WT and had a lower maximal q_{glc}, we propose that the 114-fold (at high μ) *acs* overexpression evoked a strong metabolic burden on the cells either by causing the PTA-ACS node to deplete the ATP pool (it is a futile cycle) and/or by significantly altered energy homeostasis through increased energy demands for Acs protein synthesis due to the overexpression. Both of the proposed mechanisms severely hurt the ATP balance which may trigger the cells to excrete acetate at a lower q_{glc} for energy generation purposes (Valgepea et al. 2013).

Our results confirm the hypothesis of Acs repression triggering acetate overflow (Valgepea et al. 2010) as deleting the Acs-inactivating gene *pka* resulted in postponed acetate overflow (Fig. 2b). Our proteome analysis of Δ*pka* showed lower Acs protein concentration and no major changes in central carbon metabolism compared to WT (Fig. 4). Therefore, it seems most probable that the reduced acetate overflow in the Δ*pka* strain arises from a higher pool of active Acs which enables increased acetate recycling capability in the PTA-ACS node.

The latter results indicated that acetate overflow could probably be further diminished if the level of active Acs would be slightly higher, but not too high as in Δ*pka acs*OE. Approximately 1.8-fold higher expression of Acs in Δ*pka* background was achieved by *arcA* knockout (Fig. 4), which is consistent with the increased *acs* expression of *E. coli* Δ*arcA* during exponential growth (Covert et al. 2004). Remarkably, the Δ*pka* Δ*arcA* double-knockout showed strongly reduced acetate overflow by the postponed onset of acetate excretion and 4-fold lower carbon flux to acetate compared to WT at maximal q_{glc} (Fig. 3a). The fact that Δ*pka* Δ*arcA* synthesised more CO_2 due to increased TCA cycle throughput (Fig. 3b) instead of acetate demonstrates an important point: directing higher amounts of recycled acetyl-CoA in the PTA-ACS node further into the TCA cycle is essential for diminishing acetate overflow. Positive effects of *arcA* knockout on TCA cycle gene expression (Covert et al. 2004; Vemuri et al. 2006) and flux throughput (Perrenoud and Sauer 2005) are expected and were also confirmed on the translational level by our proteome analysis (Fig. 4). It is important to note that the Δ*pka* Δ*arcA* strain would enable the production of target compounds in the absence of acetate at considerably higher values of q_{glc}, most probably resulting in higher volumetric productivities (~22 % higher gDCW L^{-1} h^{-1} compared to WT). Moreover, this strain accumulated no other detrimental by-product and showed identical μ_{max} and only ~5 % lower Y_{xs} compared to WT. This work demonstrates that a simple genetic overexpression does not work in all cases for achieving the desired effects, but the expression level of the relevant fraction of the protein pool (active Acs in this case) has to be fine-tuned together with downstream throughput (TCA). To the best of our knowledge, this is the first successful application of the modification of protein acetylation for metabolic engineering in *E. coli*.

Several explanations for acetate overflow propose limitations in respiratory capacity (Han et al. 1992; Paalme et al.

1997; Varma and Palsson 1994), TCA cycle throughput (Majewski and Domach 1990; Veit et al. 2007) or activity of the GS (Shin et al. 2009; Waegeman et al. 2011). Our results suggest that none of these pathways is limiting and triggering acetate overflow in WT *E. coli* per se since we detected lower expression of cytochrome and GS proteins in Δ*pka* Δ*arcA* compared to WT and higher acetate excretion in Δ*arcA* with up-regulated TCA cycle protein expression compared to Δ*pka* Δ*arcA* and WT (Figs. 3a and 4). More recently, Zhuang et al. (2011) proposed that acetate overflow could result from the competition for membrane space between respiratory chain enzymes and glucose transporters. However, this hypothesis has been questioned (Huberts et al. 2012), and also our proteomic data of this (Fig. 4) and of a previous study (Valgepea et al. 2010) are not consistent with Zhuang et al. (2011) as expression of respiratory chain and TCA cycle enzymes showed different behaviour. Hence, we conclude that a fine-tuned coordination between increasing the recycling capabilities of acetate in the PTA-ACS node through a higher concentration of active acetate scavenging Acs protein and downstream metabolism throughput in the TCA cycle is necessary for diminishing overflow metabolism of acetate in *E. coli*. We see this work being a good example for proving the value of systems biology study of metabolism since the strains with potential interest to industrial use reported in this work were created based on the results of previous system-wide studies of *E. coli* metabolism (Valgepea et al. 2010, 2011, 2013).

Acknowledgments The authors thank Liisa Arike and Andres Maser for their help with absolute protein quantification and fermentation experiments, respectively. The financial support for this work was provided by the European Regional Development Fund project EU29994; Ministry of Education, Estonia, through the grant SF0140090s08; and Estonian Science Foundation through grants G8165 and G9192. The data deposition to the ProteomeXchange Consortium was supported by the PRIDE team.

References

Abdel-Hamid AM, Attwood MM, Guest JR (2001) Pyruvate oxidase contributes to the aerobic growth efficiency of *Escherichia coli*. Microbiology 147:1483–1498

Adamberg K, Seiman A, Vilu R (2012) Increased biomass yield of *Lactococcus lactis* by reduced overconsumption of amino acids and increased catalytic activities of enzymes. PLoS ONE 7(10): e48223

Anfora AT, Halladin DK, Haugen BJ, Welch RA (2008) Uropathogenic *Escherichia coli* CFT073 is adapted to acetatogenic growth but does not require acetate during murine urinary tract infection. Infect Immun 76(12):5760–5767

Arike L, Valgepea K, Peil L, Nahku R, Adamberg K, Vilu R (2012) Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. J Proteome 75(17):5437–5448

Asara JM, Christofk HR, Freimark LM, Cantley LC (2008) A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. Proteomics 8(5):994–999

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008

Baldwin WW, Myer R, Kung T, Anderson E, Koch AL (1995) Growth and buoyant density of *Escherichia coli* at very low osmolarities. J Bacteriol 177(1):235–237

Barak R, Abouhamad WN, Eisenbach M (1998) Both acetate kinase and acetyl coenzyme A synthetase are involved in acetate-stimulated change in the direction of flagellar rotation in *Escherichia coli*. J Bacteriol 180(4):985–988

Barak R, Prasad K, Shainskaya A, Wolfe AJ, Eisenbach M (2004) Acetylation of the chemotaxis response regulator CheY by acetyl-CoA synthetase purified from *Escherichia coli*. J Mol Biol 342(2): 383–401

Barker CS, Prüss BM, Matsumura P (2004) Increased motility of *Escherichia coli* by insertion sequence element integration into the regulatory region of the *flhD* operon. J Bacteriol 186(22):7529–7537

Castaño-Cerezo S, Pastor JM, Renilla S, Bernal V, Iborra JL, Cánovas M (2009) An insight into the role of phosphotransacetylase (*pta*) and the acetate/acetyl-CoA node in *Escherichia coli*. Microb Cell Factories 8:54

Castaño-Cerezo S, Bernal V, Blanco-Catalá J, Iborra JL, Cánovas M (2011) cAMP-CRP co-ordinates the expression of the protein acetylation pathway with central metabolism in *Escherichia coli*. Mol Microbiol 82(5):1110–1128

Cayley S, Lewis BA, Guttman HJ, Record MT Jr (1991) Characterization of the cytoplasm of *Escherichia coli* K-12 as a function of external osmolarity. Implications for protein-DNA interactions in vivo. J Mol Biol 222(2):281–300

Contiero J, Beatty CM, Kumari S, DeSanti CL, Strohl WR, Wolfe AJ (2000) Effects of mutations in acetate metabolism in high-cell-density growth of *Escherichia coli*. J Ind Microbiol Biotechnol 24(6):421–430

Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BØ (2004) Integrating high-throughput and computational data elucidates bacterial networks. Nature 429(6987):92–96

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26(12):1367–1372

Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97(12):6640–6645

De Mey M, Lequeux GJ, Beauprez JJ, Maertens J, Van Horen E, Soetaert WK, Vanrolleghem PA, Vandamme EJ (2007) Comparison of different strategies to reduce acetate formation in *Escherichia coli*. Biotechnol Prog 23(5):1053–1063

Dittrich CR, Bennett GN, San KY (2005) Characterization of the acetate-producing pathways in *Escherichia coli*. Biotechnol Prog 21(4): 1062–1067

El-Mansi EM, Holms WH (1989) Control of carbon flux to acetate excretion during growth of *Escherichia coli* in batch and continuous cultures. J Gen Microbiol 135(11):2875–2883

Görke B, Stülke J (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat Rev Microbiol 6(8): 613–624

Han K, Lim HC, Hong J (1992) Acetic acid formation in *Escherichia coli* fermentation. Biotechnol Bioeng 39(6):663–671

Hendrickson EL, Xia Q, Wang T, Leigh JA, Hackett M (2006) Comparison of spectral counting and metabolic stable isotope labeling for use with quantitative microbial proteomics. Analyst 131(12): 1335–1341

Huberts DH, Niebel B, Heinemann M (2012) A flux-sensing mechanism could regulate the switch between respiration and fermentation. FEMS Yeast Res 12(2):118–128

Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, Ho PY, 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 (2007) Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. Science 316(5824):593–597

Kasemets K, Drews M, Nisamedtinov I, Adamberg K, Paalme T (2003) Modification of A-stat for the characterization of microorganisms. J Microbiol Methods 55(1):187–200

Klein AH, Shulla A, Reimann SA, Keating DH, Wolfe AJ (2007) The intracellular concentration of acetyl phosphate in *Escherichia coli* is sufficient for direct phosphorylation of two-component response regulators. J Bacteriol 189(15):5574–5581

Lahtvee P-J, Adamberg K, Arike L, Nahku R, Aller K, Vilu R (2011) Multi-omics approach to study the growth efficiency and amino acid metabolism in *Lactococcus lactis* at various specific growth rates. Microb Cell Factories 10:12

Lin H, Castro NM, Bennett GN, San K-Y (2006) Acetyl-CoA synthetase overexpression in *Escherichia coli* demonstrates more efficient acetate assimilation and lower acetate accumulation: a potential tool in metabolic engineering. Appl Microbiol Biotechnol 71(6): 870–874

Loferer-Krössbacher M, Klima J, Psenner R (1998) Determination of bacterial cell dry mass by transmission electron microscopy and densitometric image analysis. Appl Environ Microbiol 64(2):688–694

Majewski RA, Domach MM (1990) Simple constrained-optimization view of acetate overflow in *E. coli*. Biotechnol Bioeng 35(7):732–738

Mayover TL, Halkides CJ, Stewart RC (1999) Kinetic characterization of CheY phosphorylation reactions: comparison of P-CheA and small-molecule phosphodonors. Biochemistry 38(8):2259–2271

Mazumdar S, Clomburg JM, Gonzalez R (2010) *Escherichia coli* strains engineered for homofermentative production of D-lactic acid from glycerol. Appl Environ Microbiol 76(13):4327–4336

Milo R (2013) What is the total number of protein molecules per cell volume? A call to rethink some published values. BioEssays 35(12): 1050–1055

Mizrahi I, Biran D, Ron EZ (2009) Involvement of the Pta-AckA pathway in protein folding and aggregation. Res Microbiol 160(1):80–84

Nahku R (2012) Validation of critical factors for the quantitative characterization of bacterial physiology in accelerostat cultures. Dissertation, Tallinn University of Technology, Tallinn

Nahku R, Valgepea K, Lahtvee P-J, Erm S, Abner K, Adamberg K, Vilu R (2010) Specific growth rate dependent transcriptome profiling of *Escherichia coli* K12 MG1655 in accelerostat cultures. J Biotechnol 145(1):60–65

Nahku R, Peebo K, Valgepea K, Barrick JE, Adamberg K, Vilu R (2011) Stock culture heterogeneity rather than new mutational variation complicates short-term cell physiology studies of *Escherichia coli* K-12 MG1655 in continuous culture. Microbiology 157:2604–2610

Narang A (2009) Quantitative effect and regulatory function of cyclic adenosine 5′-phosphate in *Escherichia coli*. J Biosci 34(3):445–463

Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1(5):376–386

Paalme T, Kahru A, Elken R, Vanatalu K, Tiisma K, Vilu R (1995) The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate (A-stat). J Microbiol Methods 24(2):145–153

Paalme T, Elken R, Kahru A, Vanatalu K, Vilu R (1997) The growth rate control in *Escherichia coli* at near to maximum growth rates: the A-stat approach. Antonie Van Leeuwenhoek 71(3):217–230

Perrenoud A, Sauer U (2005) Impact of global transcriptional regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on glucose catabolism in *Escherichia coli*. J Bacteriol 187(9):3171–3179

Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc 2(8):1896–1906

Renilla S, Bernal V, Fuhrer T, Castaño-Cerezo S, Pastor JM, Iborra JL, Sauer U, Cánovas M (2011) Acetate scavenging activity in *Escherichia coli*: interplay of acetyl-CoA synthetase and the PEP-glyoxylate cycle in chemostat cultures. Appl Microbiol Biotechnol 93(5):2109–2024

Sharan SK, Thomason LC, Kuznetsov SG, Court DL (2009) Recombineering: a homologous recombination-based method of genetic engineering. Nat Protoc 4(2):206–223

Shiloach J, Reshamwala S, Noronha SB, Negrete A (2010) Analyzing metabolic variations in different bacterial strains, historical perspectives and current trends—example *E. coli*. Curr Opin Biotechnol 21(1):21–26

Shin S, Chang DE, Pan JG (2009) Acetate consumption activity directly determines the level of acetate accumulation during *Escherichia coli* W3110 growth. J Microbiol Biotechnol 19(10):1127–1134

Starai VJ, Escalante-Semerena JC (2004) Identification of the protein acetyltransferase (Pat) enzyme that acetylates acetyl-CoA synthetase in *Salmonella enterica*. J Mol Biol 340(5):1005–1012

Starai VJ, Celic I, Cole RN, Boeke JD, Escalante-Semerena JC (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science 298(5602):2390–2392

Taymaz-Nikerel H, De Mey M, Baart G, Maertens J, Heijnen JJ, van Gulik W (2013) Changes in substrate availability in *Escherichia coli* lead to rapid metabolite, flux and growth rate responses. Metab Eng 16:115–129

Valgepea K, Adamberg K, Nahku R, Lahtvee P-J, Arike L, Vilu R (2010) Systems biology approach reveals that overflow metabolism of acetate in *Escherichia coli* is triggered by carbon catabolite repression of acetyl-CoA synthetase. BMC Syst Biol 4:166

Valgepea K, Adamberg K, Vilu R (2011) Decrease of energy spilling in *Escherichia coli* continuous cultures with rising specific growth rate and carbon wasting. BMC Syst Biol 5:106

Valgepea K, Adamberg K, Seiman A, Vilu R (2013) *Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins. Mol BioSyst 9(9):2344–2358

Varma A, Palsson BØ (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. Appl Environ Microbiol 60(10):3724–3731

Veit A, Polen T, Wendisch VF (2007) Global gene expression analysis of glucose overflow metabolism in *Escherichia coli* and reduction of aerobic acetate formation. Appl Microbiol Biotechnol 74(2):406–421

Vemuri GN, Altman E, Sangurdekar DP, Khodursky AB, Eiteman MA (2006) Overflow metabolism in *Escherichia coli* during steady-state growth: transcriptional regulation and effect of the redox ratio. Appl Environ Microbiol 72(5):3653–3661

Vizcaíno JA, Côté RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J, O'Kelly G, Schoenegger A, Ovelleiro D, Pérez-Riverol Y, Reisinger F, Ríos D, Wang R, Hermjakob H (2013) The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. Nucleic Acids Res 41:D1063–D1069

Waegeman H, Beauprez J, Moens H, Maertens J, De Mey M, Foulquié-Moreno MR, Heijnen JJ, Charlier D, Soetaert W (2011) Effect of *iclR* and *arcA* knockouts on biomass formation and metabolic fluxes in *Escherichia coli* K12 and its implications on understanding the metabolism of *Escherichia coli* BL21 (DE3). BMC Microbiol 11:70

Weinert BT, Iesmantavicius V, Wagner SA, Schölz C, Gummesson B, Beli P, Nyström T, Choudhary C (2013) Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. Mol Cell 51(2): 265–272

Wiśniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. Nat Methods 6(5):359–362

Wiśniewski JR, Ostasiewicz P, Duś K, Zielińska DF, Gnad F, Mann M (2012) Extensive quantitative remodeling of the proteome between normal colon tissue and adenocarcinoma. Mol Syst Biol 8:611

Wolfe AJ (2005) The acetate switch. Microbiol Mol Biol Rev 69(1):12–50

Yang YT, Bennett GN, San K-Y (1999) Effect of inactivation of *nuo* and *ackA-pta* on redistribution of metabolic fluxes in *Escherichia coli*. Biotechnol Bioeng 65(3):291–297

Zhao K, Chai X, Marmorstein R (2004) Structure and substrate binding properties of cobB, a Sir2 homolog protein deacetylase from *Escherichia coli*. J Mol Biol 337(3):731–741

Zhuang K, Vemuri GN, Mahadevan R (2011) Economics of membrane occupancy and respiro-fermentation. Mol Syst Biol 7:500

PUBLICATION IV

Valgepea K, Adamberg K, Seiman A & Vilu R

***Escherichia coli* achieves faster growth by increasing catalytic and translation rates of proteins.**

Mol. Biosyst. 9, 2344–2358, (2013)[2]

[2]Copyright 2013, reprinted with permission from The Royal Society of Chemistry

Molecular BioSystems

RSC Publishing

PAPER

Escherichia coli achieves faster growth by increasing catalytic and translation rates of proteins†

Cite this: *Mol. BioSyst.*, 2013, **9**, 2344

Kaspar Valgepea,*[ab] Kaarel Adamberg,[abc] Andrus Seiman[b] and Raivo Vilu[ab]

Regulation levels of the gene expression cascade controlling protein levels and metabolic fluxes for cells to achieve faster growth have not been elaborated in acceptable detail. Furthermore, there is need for specific growth rate (μ) dependent absolute quantitative transcriptome and proteome data to understand the molecular relationships for enabling cells to modify μ. We address these questions, for the first time, by presenting regulatory strategies for more efficient metabolism of *Escherichia coli* at higher μ by statistical covariance analysis of genome-wide intracellular mRNA and protein concentrations coupled to metabolic flux analysis in the steady state range of μ = 0.11–0.49 h^{-1}. Our analyses show dominating post-transcriptional control of protein abundances and post-translational control of flux rates. On average, *E. coli* achieved five-times faster growth through 3.7-fold increase of apparent catalytic rates of enzymes (k_{app}) and 2.5-fold increased translation rates, demonstrating the relevance of post-translational regulation for increasing flux throughput. Interestingly, pathways carrying the highest flux showed both high protein abundance and k_{app} values. Furthermore, co-regulation analysis of enzymatic capacities revealed tightly coupled regulatory dependencies of protein synthesis and RNA precursor synthesis, substrate utilization, biosynthetic and energy generation pathways carrying the highest flux. We also observed metabolic pathway and COG specific protein and metabolic flux control levels, protein expression costs and genome-wide principles for translation efficiency and transcription unit polarity. This work contributes to the much needed quantitative understanding of coordinated gene expression regulation and metabolic flux control. Our findings will also advance modeling and metabolic engineering of industrial strains.

Received 25th March 2013,
Accepted 31st May 2013

DOI: 10.1039/c3mb70119k

www.rsc.org/molecularbiosystems

Introduction

Most microorganisms can grow under very different environmental conditions and use a wide range of substrates, which presents a significant challenge to the flexibility of their metabolism. Adaptation to different environmental conditions is usually accompanied by a change in the specific growth rate (μ), which integrates regulation of many properties of cell physiology, *e.g.* cell size, metabolic flux patterns, energy production. The majority of these mechanisms are regulated through changes in gene expression. Gene expression regulation of enzymes and subsequent changes in metabolic flux patterns are, however, not a straightforward consequence of transcriptional regulation as once thought.[1] The gene expression cascade is under multi-level regulation: transcription can be regulated through molecular modifications of DNA (*e.g.* methylation); binding of transcription factors; regulation of translation through mechanisms directly interacting with the ribosome or its associated initiation proteins; mRNA and protein degradation; post-translational chemical modification of proteins (*e.g.* phosphorylation, ubiquitination), *etc*. As cells have to increase throughput of metabolic fluxes to grow faster, the question—at which regulation levels of the gene expression cascade (transcriptional (TR), translational (TL), post-TR, post-TL) is control of metabolic fluxes in different metabolic pathways realized for achieving faster growth—emerges as highly important. Control mechanisms of metabolic fluxes, *i.e. in vivo* reaction rates, are central to understanding regulation of metabolism since they represent the integrated response of all levels of cellular regulation.[2]

So far, mainly hierarchical regulation analysis has been used as the methodology for determining to which extent a change

[a] *Tallinn University of Technology, Department of Chemistry, Akadeemia tee 15, 12618 Tallinn, Estonia*
[b] *Competence Centre of Food and Fermentation Technologies, Akadeemia tee 15a, 12618 Tallinn, Estonia. E-mail: kaspar@tftak.eu; Fax: +3726408282; Tel: +3726408200*
[c] *Tallinn University of Technology, Department of Food Processing, Ehitajate tee 5, 19086 Tallinn, Estonia*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3mb70119k

in a particular flux is regulated by either gene expression or metabolism (enzyme activities).[3] Most of the studies implementing this approach have investigated yeast and concluded that fluxes through glycolytic and selected fermentative pathways are regulated in a diverse and subtle way with post-TR regulation being the dominant mechanism (reviewed in van Eunen *et al.*[4]). A notable study using this approach is that of glycolysis metabolism perturbation by the absence of oxygen and the presence of ATP-depleting benzoic acid compared to aerobic conditions in *Saccharomyces cerevisiae*, where the authors were able to dissect gene expression regulation in glycolysis into five layers: TR, TL, post-TL, mRNA and protein degradation.[5] However, these and other studies in the literature, to the best of our knowledge, have not been carried out to study global μ-dependent protein abundance and flux control levels, thus not producing genome-wide quantitatively reliable data which could be used for the elucidation of the principles of growth regulation and quantitative modeling. Hence, a more comprehensive approach is needed.

Although many μ-dependent cellular global relationships have been determined in *E. coli* batch cultures,[6] high-quality experimentally determined quantitative interdependencies of μ and the most relevant cellular characteristics including transcriptome,[7–9] metabolome and fluxome,[8,10,11] proteome,[12,13] biomass composition,[11,12] and cell volume[14,15] are scarce. What is more, μ-dependent genome-wide absolute quantitative mRNA and proteome data (molecules per cell and concentrations) for *E. coli* are missing altogether. A comprehensive understanding of the physiological processes in cells calls for a systems biology effort of integrating high-quality absolute quantitative transcriptomic, proteomic and metabolomic data coupled to models of different levels of detail.[16–18] This may seem as an immense challenge, but as understanding of both the regulation levels of metabolic fluxes and other molecular relationships enabling the cells to modify μ is of instrumental importance towards a more complete description of the control principles of cell metabolism[16,19] and successful biotechnology and synthetic biology efforts,[20] genome-wide absolute quantitative multi-level analysis in one study is highly needed. Moreover, it is important that this biological information is acquired from cell cultures in strictly defined physiological states, *i.e.* continuous cultures.[21,22] Exactly this has been achieved in the present paper.

We previously carried out triplicate μ-dependent glucose-limited continuous culture (accelerostat (A-stat)[23]) experiments with *E. coli* K-12 MG1655 coupled to relative mRNA and protein expression and metabolic flux analyses.[11,13] These experiments allowed us to propose a new theory for overflow metabolism of acetate[13] and detect decreased energy spilling with rising μ and carbon wasting.[11] To understand the observed phenomena and regulation of metabolism further we performed in this study an absolute quantitative *in vivo* biochemistry analysis; *i.e.* counting of mRNAs and proteins per cell and measuring their intracellular concentrations; characterization of translational efficiencies by protein-per-mRNA ratios; determination of protein and flux control levels by statistical covariance analysis; estimation of *in vivo* catalytic rates of enzymes, *etc*. More specifically, the aim of this work was to determine genome-wide μ-dependent gene expression regulation levels for controlling protein abundances and flux rates in *E. coli* for achieving faster growth. Flux changes were evoked by the A-stat method,[23] which enables study of the cells in strictly defined physiological states determined by controlled physico-chemical conditions and continuous monitoring of the dynamic adaptation of cells under the conditions of changing μ. Determination of protein abundance and flux control levels, and quantitative description of various molecular, energetic and metabolic aspects of the gene expression cascade and different metabolic pathways was realized through integration of absolute transcriptome and proteome data with calculated *in vivo* flux rates published previously[11] in the range of μ = 0.11–0.49 h^{-1}.

Results and discussion

Global μ-dependent absolute proteome and transcriptome

We determined both the intracellular abundance (molecules per cell) and concentration (molecules per fL of biomass) for all the quantified mRNA and protein molecules taking into account also the change in cell volume, and total RNA and protein concentration in the biomass with increasing μ (Table S1, ESI†). Genome-wide absolute quantitative transcriptome data for μ = 0.11; 0.21; 0.31; 0.40; 0.48 h^{-1} were determined using DNA microarrays for around 4300 mRNAs. Coefficient of variation (CV) between six DNA microarray technical replicates at μ = 0.11 h^{-1} was 11%. Absolute proteome quantification was performed for 1185 proteins at μ = 0.11 h^{-1} using the iBAQ approach[24] and published in Arike *et al.*[25] Pearson correlation coefficient (PCC) and CV between two independent A-stat experiments were R = 0.99 and 11%, respectively.[25] Protein abundances for μ = 0.20; 0.30; 0.40; 0.49 h^{-1} were calculated based on relative data published previously in Valgepea *et al.*[13] which also showed high correlation between independent experiments (R = 0.79–0.92 and CV = 11%). High confidence of our absolute quantitative proteome data set is indicated by the fact that the sum of all quantified proteins by iBAQ was on average only 10% lower than the cellular total protein concentration determined by Lowry analysis at all μ (Table S1, ESI†). The complete gene-specific data set of absolute and relative mRNA, protein and specific flux values, covariance values and other parameters used in calculations can be found in Table S2 (ESI†).

The proteomic coverage of protein-coding quantified mRNAs reached 100% with higher mRNA concentrations (Fig. 1A) being in line with the proposed 'lazy step function' (*i.e.* the ability to detect proteins rising at higher mRNA levels) of protein identification in *E. coli*, *Bacillus subtilis* and higher organisms.[26,27] Both mRNA and protein concentrations spanned approximately three orders of magnitude (Fig. 1B) while protein concentrations were on average 1000-fold higher than corresponding transcripts. A larger dynamic range of protein abundances has been observed in yeast[28] and mammalian cells.[27] Our genome-wide mRNA and protein concentration data show

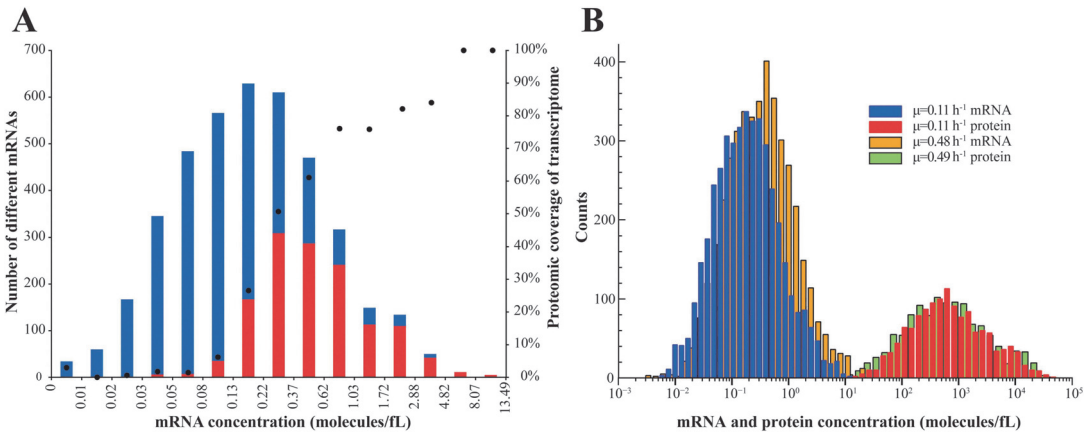

Fig. 1 Global absolute proteome and transcriptome characteristics. (A) Proteomic coverage of protein-coding quantified mRNAs at μ = 0.11 h^{-1}. All quantified protein-coding mRNAs are binned by their concentration. The number of different mRNAs and mRNAs with respectively quantified translation product—protein—belonging to a particular bin is indicated in blue and red, respectively. Black circles denote the proteomic coverage of transcriptome. (B) mRNA and protein concentration dynamic range at low and high μ.

significant and increasing correlations of R = 0.62–0.78 with rising μ (Fig. S1, ESI†).

As expected with increasing cell size, the sum of mRNAs and proteins per cell increased 4.2- and 2-fold, respectively, from μ = 0.11 to 0.49 h^{-1} (Fig. S2A, ESI†). This is in concordance with the increase of the RNA-to-protein mass ratio with rising μ also observed previously in *E. coli*.[29,30] The sum of intracellular mRNA and protein concentrations (molecules per fL) showed a different behavior: 1.8-fold increase for mRNAs and 1.2-fold decrease for proteins was observed (Fig. S2B, ESI†). Similarly, opposite μ-dependence of protein abundances and concentrations has also been observed for unregulated constitutive genes in *E. coli*.[31] This indicates a faster increase of cell volume compared to protein abundance with faster growth. This different μ-dependent mRNA and protein abundance and concentration behavior is an important physiological observation and essential for taking into account in *in silico* modeling approaches where cell metabolism is simulated as a function of μ.[31,32]

Majority of protein levels are controlled at the post-TR level

To address the main question of this study—at which regulation levels of gene expression does *E. coli* control its protein levels and metabolic fluxes for achieving faster growth—we integrated our quantitative proteome and transcriptome data with previously published flux measurements from the same experiments.[11] We used covariance analysis to statistically determine the regulation levels of gene expression—TR, post-TR, TL, post-TL—as it describes both the direction and magnitude of mRNA, protein and flux changes with increasing μ making it a suitable statistical method for analysis of absolute quantitative data (see Experimental).

Firstly, we analyzed the impact of TR, post-TR and TL regulation on protein abundance levels for *E. coli* to achieve faster growth using the protein-per-mRNA ratio (pm). This ratio estimates translation efficiency and the changes give indications about the level of gene expression regulation either through protein translation or degradation making it an important molecular parameter. Values of pm for ~1200 genes ranged from around 100 to 10 000 at low μ compared to 50 to 4000 at high μ (Fig. S3, ESI†) pointing to different levels of regulation among genes through translation and protein degradation, post-translational modification and possible functional requirements for protein binding,[33–35] and explain the non-perfect mRNA and protein correlations. Plotting pm values against protein concentrations showed that abundant proteins possess about 100-times higher pm values than low abundant ones regardless of μ (Fig. S3, ESI†). It was also observed that genes belonging to clusters of orthologous groups (COG)[36] of translational machinery (J), energy generation (C) and post-translational modification (O) showed higher pm ratios (Fig. S4, ESI†) indicating the importance of efficient translation of these enzymes. Interestingly, pm values saturated at around 4000 and 2500 at low and high μ, respectively (Fig. S3, ESI†). A similar translation efficiency difference between low and high abundant proteins and the saturation effect are seen in yeast[28] and for translation rates in mammalian cells.[24] Median pm decreased 2.3-fold from 1532 to 656 with μ increasing from 0.11 to 0.48 h^{-1}. The lower ratio determined at μ = 0.48 h^{-1} is close to those of log-phase *E. coli* cells.[37,38] This is an important finding for modeling μ-dependent gene expression since constant translation rates have previously been used in simulations.[31]

TR, post-TR and TL regulation of protein abundances was quantitatively described by calculating covariance between each gene's pm and μ. The obtained covariance values were subjected to statistical hypothesis testing of being significantly (p-value < 0.05) not different or different from zero (all tests were subject to correction by false discovery rate filtering

according to the Benjamini–Hochberg procedure at level $\alpha = 0.05$). Data for all genes with covariance, *p*-values and respective mRNA and protein expression burden with rising μ can be seen in Data S1 (ESI†). A covariance value not different from zero represents TR regulation of gene expression by the pm ratio being constant at all μ. We note here that TR regulation can also be attained through changing mRNA degradation rates but the mechanism through which mRNA concentrations are controlled does not affect the calculation of pm values and the respectively derived regulation levels. Strikingly, from the total of 1112 analyzed genes with both quantified mRNA and protein concentrations, only 25% (275) showed TR regulation (Table S3, ESI†). Genes involved in glyoxylate shunt (GS), NADH metabolism and various degradation pathways showed high fraction of TR regulation (Table S4, ESI†). However, for the majority of genes (56%; 627), protein concentrations were controlled at the post-TR level determined by their covariance value being negative and different from zero (Fig. 2A), meaning decreasing pm with rising μ (protein concentrations increase less than mRNA ones or decrease more than mRNA). Particularly high enrichment of post-TR regulated genes was observed in COGs of cell cycle (D), translation (J), amino acid metabolism (E) (Table S5, ESI†) and amino acid synthesis pathways (Table S4, ESI†). Also the protein levels in the pathways carrying the highest flux—glycolysis and TCA cycle—were regulated at post-TR level for >60% of genes. We detected only four genes (0.4%) with TL level regulation indicated by positive covariance values different from zero, meaning increasing pm with rising μ (protein concentrations increase more than mRNA ones or decrease less than mRNA). This small number could be considered as an indication for the faster increase of the transcription rate compared to the translational rate with growing μ.[31]

Similar domination of post-TR regulation of protein concentrations at mRNA and protein levels was lately reported in a study of the fermentative bacterium *Lactococcus lactis*.[39] Furthermore, post-TR regulation is the predominant mechanism for controlling pm ratios in *Mycoplasma pneumoniae*,[40] mammalian cells[24] and also for ~37% of genes in *B. subtilis*.[41] Transcription was also shown to have limited control over capacities of key central metabolism enzymes in *E. coli* with faster growth in chemostat cultures.[42] In addition, several studies on yeast using hierarchical regulation analysis have also concluded that protein expression in glycolytic and fermentative pathways is mainly regulated at the post-TR level.[4,5] Thus it seems that post-TR regulation of protein levels is significant in organisms with very different growth characteristics, and practically in all the main metabolic subsystems.

Flux control is mainly exerted through post-TL regulation

Next, we moved one layer up to determine the regulation levels of metabolic fluxes for *E. coli* to achieve faster growth. Measurement of both protein concentrations and flux rates in this study enabled us to quantitatively determine the impact of change in protein concentration and its catalytic rate for realizing higher flux throughput. For this, catalytic rates of enzymes were calculated as the ratio of specific flux and protein concentration and defined as the apparent catalytic rate (k_{app}, s^{-1}). Flux control levels—TL or post-TL—were determined by covariance analysis of changing k_{app} and μ. Post-TL regulation includes modifications of proteins after translation (post-TL chemical modification or allosteric regulation) or change of protein activity due to hyperbolic change of enzyme kinetics solely due to substrate concentration changes.[43] As, to the best of our knowledge, no reliable intracellular metabolome quantification method yet exists for *E. coli* continuous cultures, it was not possible to experimentally investigate the latter possibility and differentiate between the two types of regulation mechanisms of enzymatic activities.

Covariance value being statistically not different from zero represents protein control of flux (TL regulation) as protein concentration increases proportionally to the specific flux rate with rising μ (*i.e.* constant k_{app}). Only 9% (18) out of 191 genes under analysis showed flux control through protein levels (TL regulation) (Table S6; Data S2, ESI†) which included genes from glutamine and nucleotide synthesis pathways (Table S4, ESI†). Notably, for 10 out of these 18 genes flux control is at the TR level as both of their pm and k_{app} values were not different from zero. Recently, transcriptional control of fluxes was shown in the TCA cycle, GS and acetate excretion but not in PPP using transcription factor mutant strains of *E. coli*.[44] The 10 genes for which we detected TR level control of fluxes did not, however, belong to any of these pathways. Flux throughput was controlled at the post-TL level for the great majority of genes (81%; 154) shown by positive covariance values different from zero, meaning increasing k_{app} with rising μ (Fig. 2B). For nearly all the genes organized into COGs of energy production and conversion (C), translation (J) and high flux pathways—TCA cycle and glycolysis—flux control was achieved through post-TL regulation (Tables S4 and S5, ESI†). These are highly abundant proteins (over 5000 copies per cell) indicating, similarly to protein abundance control (see above), that *E. coli* has to implement additional regulatory processes for increasing their enzymatic capacities to fulfill the demands of rising biomass and energy synthesis throughput with faster growth. This is also reflected by the observation that the average protein abundance and synthesis cost for post-TL compared to TL-regulated genes were ~2-fold higher (5190 *vs.* 2923 molecules per cell and 6.2×10^6 *vs.* 3.1×10^6 ATP molecules per cell, respectively). Thus it seems to be energetically favorable for *E. coli* to increase the catalytic capacity of abundant proteins through post-TL processes and save ATP from lower protein synthesis costs. These results are in contrast to an *in silico* study which proposed that low-cost enzymes in *E. coli* are less likely to be post-TR regulated.[45]

Predominant post-TL control of several central metabolism fluxes in *E. coli* has been reported using *in vitro* enzyme assays when cells were shifted from low to high dilution rates in chemostats.[42] Our results are also in accord with an *E. coli* study based on EcoCyc database[46] information on gene expression regulation, which showed coupling of energy generating reactions to enzymatic regulation, important for short-term maintenance of energy homeostasis.[47] Furthermore, prevalence

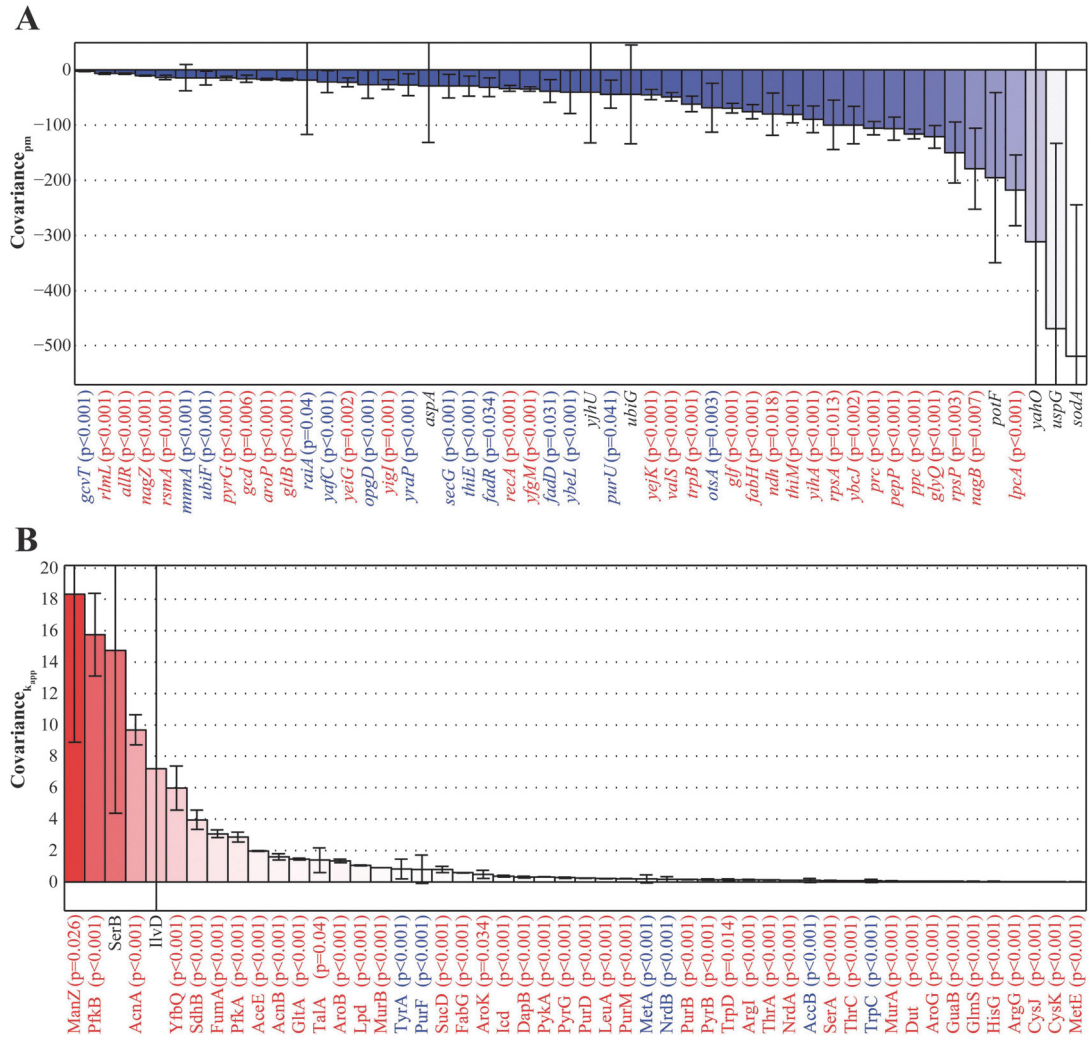

Fig. 2 Covariance analysis of protein concentration and flux control levels. (A) Covariance analysis of control of protein concentrations for 50 randomly chosen genes. Covariance values are calculated between each gene's pm values and μ. Red and blue colored bars indicate increasing and decreasing pm with rising μ, respectively. Colors of gene names: red denotes covariance value being statistically significantly different from zero, meaning post-TR control of protein concentration; blue denotes covariance value being statistically significantly not different from zero, meaning TR control of protein concentration; black denotes genes with too large measurement error for regulation determination. Data for all genes can be seen in Data S1 (ESI†). Error bars denote 95% confidence intervals of covariance values. Error bars have been calculated using error propagation principles and individual errors of mRNA and protein concentrations. *p*-values represent the results of the statistical hypothesis testing of covariance values being statistically significantly different or non-different from zero. Refer to Experimental for detailed description of covariance calculation, error analysis and statistical significance testing. (B) Covariance analysis of flux control for 50 randomly chosen genes. Covariance values are calculated between each gene's k_{app} values and μ. Red colored bars indicate increasing k_{app} with rising μ. Colors of gene names: red denotes covariance value being statistically significantly different from zero, meaning post-TL flux control; blue denotes covariance value being statistically significantly not different from zero, meaning TL flux control; black denotes genes with too large measurement error for regulation determination. Data for all genes can be seen in Data S2 (ESI†). Refer to Fig. 2A legend for description of error bars and statistical analysis.

of post-TL regulation in controlling central metabolism flux throughput was also observed in *L. lactis*,[39] and has been recently concluded to be probably the primary flux controlling mechanism based on numerous studies.[18] As a substantial amount of quantitative evidence for the importance of post-TL regulation in flux control has now accumulated, it will be highly relevant in the future to further dissect the specific post-TL regulation mechanisms in action when a considerable amount of information on protein specific post-TL modifications and allosteric regulation by reactant metabolites becomes available.[18,43,48]

Apparent catalytic rates (k_{app}) of central metabolism enzymes

E. coli had to increase its specific flux rates ~5-fold for achieving faster growth in the studied range of μ = 0.11–0.49 h^{-1}. Higher flux throughput can be realized by increasing enzymatic capacities either through increased protein abundances or catalytic rates. Fluxes in *E. coli* were predominantly controlled through post-TL regulation as presented above. Indeed, we detected a median 3.7-fold increase of k_{app} of 191 central metabolism and biosynthetic enzymes (Fig. S5, ESI†) compared to non-changing median protein concentration with rising μ confirming that higher flux throughput for faster growth in *E. coli* is mainly achieved through increased catalytic rates of enzymes. The same principle for flux control also applies to the most important pathways for aerobic growth of *E. coli*: the average protein concentrations of glycolysis, the TCA cycle, the pentose phosphate pathway (PPP) and acetate synthesis were maintained constant or even decreased with rising μ and higher flux throughput was clearly realized through increasing enzyme k_{app} (Fig. 3). This further supports the conclusion of the relevance of post-TL regulation in metabolic control as these are the pathways carrying the highest flux and responsible for feeding precursor molecules into energy generation and biosynthetic pathways. The rising k_{app} values possibly contribute to the increased metabolic efficiency for enabling *E. coli* to maintain a constant biomass yield under increased carbon wasting with rising μ.[11] Our results are in line with a similar recent report for *L. lactis* where 3.6-fold increase of average k_{app} for central metabolic and biosynthetic enzymes with 5-fold faster growth contributed to rising biomass yield.[39]

When comparing the actual values of k_{app} among the most important pathways, one can see that the average values for the TCA cycle (49–156 s^{-1} from μ = 0.11 to 0.49 h^{-1}, respectively) are higher than for glycolysis (34–110 s^{-1}) and PPP (4–88 s^{-1}) (Fig. 3) indicating that pathways with higher contribution of ATP production to the total ATP pool also possess enzymes with higher catalytic rates. Notably, these k_{app} values closely match the range of k_{cat} values for central carbon metabolism enzymes in *E. coli* measured *in vitro*.[49] The same can be observed at the COG functional class level as k_{app} values of energy metabolism enzymes (C) were an order or two higher than those of biomass monomer synthesis (E, F, H, I) (Fig. S6, ESI†). Similar higher (~30-fold) catalytic rates for central compared to secondary metabolism enzymes is seen when ~2000 k_{cat} values measured *in vitro* for prokaryotes and eukaryotes were analyzed.[49] Hence, cells maintain higher abundance of proteins required for biomass synthesis (translational machinery, monomer precursor synthesis) whereas enzymatic capacities for energy generating proteins are more likely to be increased through post-TL regulation. Interestingly, we also noted that the pathways carrying the highest flux—glycolysis, TCA cycle and PPP—showed both higher protein concentration and k_{app} compared to biosynthetic pathways (Fig. 4). This refers to an evolutionary push towards proteins carrying high flux being more abundant and possessing higher catalytic rates to reduce the cost of protein production, as also proposed previously.[49] Overall, it seems that under strong nutrient limitation at low μ, metabolism of *E. coli* is on 'standby' mode: protein abundances are high and catalytic rates not

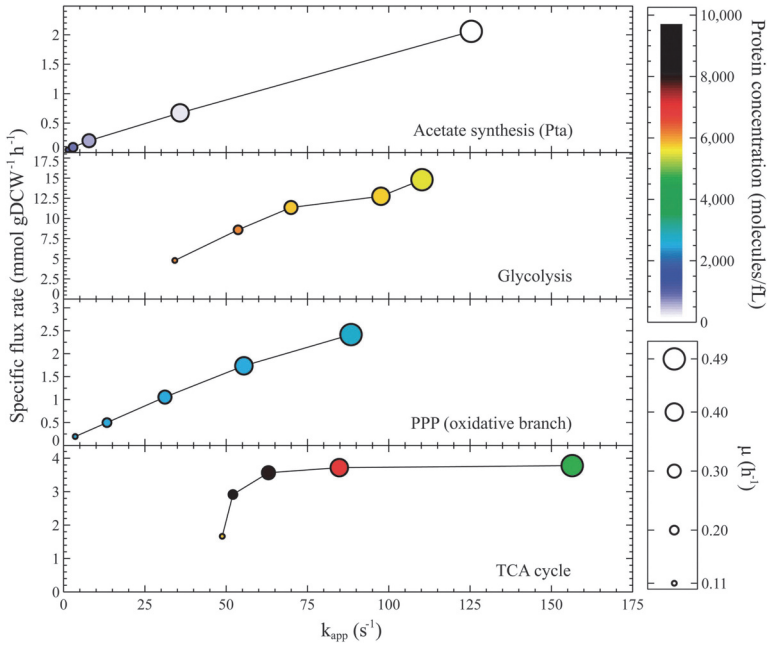

Fig. 3 Flux control of the most important pathways for aerobic growth of *E. coli*. Each circle represents the average value of the pathway if not otherwise noted. See Table S7 (ESI†) for the genes assigned to pathways according to the EcoCyc database.[46] gDCW, gram of dry cell weight.

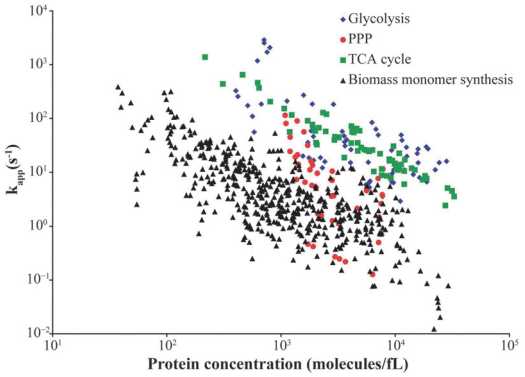

Fig. 4 Protein concentration and their k_{app} values for pathways carrying the highest flux compared to biosynthetic pathways. See Table S7 (ESI†) for the genes assigned to pathways according to the EcoCyc database.[46]

saturated for biomass and energy generation so that cells could quickly respond to changing environmental conditions by modifying catalytic rates of proteins without wasting time for increasing their levels. An analogy for this could be drawn from drag racing: drivers 'pump up' the revolutions of the engines of the cars before the start signal so they could instantly take off with full torque once the light turns green. The same phenomenon is also demonstrated by glucose-pulse experiments where cells are able to rapidly increase μ 3-fold after a substrate pulse,[50,51] clearly pointing to the ability to immediately increase catalytic rates if needed.

Our genome-wide quantitative data enable us to investigate patterns of protein catalytic rates within many pathways. In *L. lactis*, first enzymes in central pathways tend to have lower abundances and, hence, higher k_{app} values,[39] also seen indirectly in yeast.[5] This hints for possible allosteric control of feedback regulation through the product of the pathway to precisely control the flux.[45] Our data are in accordance with the latter and provide further pathway-specific observations. Both initial enzymes of glycolysis and PPP, Pgi and Zwf, respectively, possess higher k_{app} values compared to subsequent enzymes probably for enabling strict distribution of carbon flow between these pathways for fast switching between the need for more ATP (glycolysis) or NADPH (PPP). One would expect that a biosynthesis pathway is activated when its product is needed and flux is controlled by its first enzyme. However, a difference among biosynthetic pathways was observed: longer pathways (*e.g.* purine, aromatic amino acids, Arg, Lys) showed higher k_{app} values for their first enzymes which was not seen for shorter ones (*e.g.* Ser, Thr, Leu, His). It might be useful for the cell to control flux through initial enzymes in longer pathways to realize a fast response as opposed to a time-consuming situation where all the enzyme levels are maintained low and increased uniformly once the pathway capacity has to be enhanced.

Ribosomal translation rate and regulation of transcription

An important parameter that cells have to adjust with adaptation to changing environmental conditions is the protein synthesis rate. For example, faster growing cells are bigger since they need more 'catalytic units' to process nutrients and synthesize biomass faster. For this, a cell needs to increase its translational capacity with rising μ for maintaining the necessary concentrations of enzymes, the cellular 'workhorses'. How is this achieved? In our experiments, the median ribosomal protein concentration in *E. coli* increased only 1.5-fold with five times faster growth (Fig. S7A, ESI†) indicating the need for higher translation rates (k_{cat} of ribosomes) to maintain sufficient translational capacity. Indeed, calculating k_{app} for ribosomal proteins, which is a good estimate for the translation rate, revealed 2.8-times (median; CV = 11% among 52 quantified ribosomal proteins) faster ribosomal translation rates with increasing μ, clearly demonstrating higher translational capacity in faster growing cells (Fig. S7B, ESI†). This experimental evidence is of particular importance since it confirms a previously postulated hypothesis that translational capacity is rate-limiting for faster growth of *E. coli*.[30] Hence, our results show that increasing the concentration of ribosomal proteins (equivalent to ribosomes) is not sufficient for achieving faster growth and, therefore, also translation rates and k_{app} values of metabolic enzymes have to be raised. The observation of increased translation rates in faster growing cells is in concordance with the literature using indirectly calculated values.[6,29,52]

Another crucial parameter for modifying concentrations of both mRNAs and proteins is the transcription rate of RNA polymerase (RNAP). The same principle as with the translation rate applies here: faster growing cells need to increase their transcriptional capacity which can be achieved through increasing RNAP concentration and/or its synthesis rate. In the literature, varying abundances from 2000 to 5000 RNAP molecules per genome equivalent exist.[53] Abundances of the RNAP subunit proteins β and β′ determined in this study fall within this range with per cell values of 2577 and 2230 at μ = 0.11 h^{-1} and increasing to 6576 and 5559 at μ = 0.49 h^{-1}, respectively (Table S2, ESI†). As the sum of mRNA abundances increased over 4-times with 5-fold faster growth (Fig. S2A, ESI†), it becomes obvious that the ~2.5-fold rise in RNAP abundance is not sufficient to meet the demands for faster RNA synthesis with increasing μ. Indeed, calculation of the RNAP synthesis rate (see Experimental) revealed that transcription speed in *E. coli* is increased ~8-fold to enable five times faster growth (μ = 0.49 *vs.* 0.11 h^{-1}). This implies a very extensive regulation of RNAP, possibly through σ-factors and the alarmone ppGpp or other proteins.

Co-regulation of protein synthesis and metabolic pathways

Cells have to successfully couple substrate utilization, biomass monomer synthesis, energy generation *etc.* to protein synthesis in order to realize the doubling of cell mass during the cell cycle. After looking separately at the regulation of enzymatic capacities over pathways and approximate translation rate for ribosomes, we next studied the co-regulation principles of metabolic pathways and protein synthesis. Since we observed that patterns in changes of k_{app} with rising μ (Fig. 3; Fig. S7B, ESI†), we applied correlation analysis of all genes' relative k_{app}

curves (k_{app} at each μ compared to reference μ = 0.11 h^{-1}) to understand which pathways work jointly and are coupled to translation.

One can see clearly that enzymatic capacities of ribosomal proteins are co-regulated with RNA precursor synthesis—pyrimidine and purine—pathways (red to green squares in Fig. 5) since k_{app} of proteins in these pathways and for ribosomal proteins changed similarly with increasing μ. Interestingly, both latter pathways contain genes whose k_{app} values increased more than k_{app} values of ribosomal proteins (green). Also the majority of amino acid synthesis pathway proteins were co-regulated with translation (*e.g.* lysine, threonine, cysteine) while proline and alanine synthesis pathways showed even faster increase of k_{app} compared to ribosomal proteins. These co-regulations are expected since RNA and amino acid synthesis have to work coherently with translation for realizing faster synthesis of proteins and cell mass with rising μ.

Also the enzymatic capacities in substrate utilization, biosynthetic and energy generation pathways carrying the highest flux—glycolysis and the TCA cycle—are tightly coupled (red to green) to protein synthesis (Fig. 5) *i.e.* change in the concentrations of intermediates of these pathways is coordinated. An interesting case is PPP which is also co-regulated with protein synthesis since all its proteins' k_{app} values increase substantially faster compared to ribosomes (green). Actually, PPP k_{app} values increased more than for almost all other proteins (green area of PPP in Data S3, ESI†). The details behind these observations can be related to NADPH metabolism and will be investigated in further studies.

Correlation analysis can also reveal the uniformity of protein regulation within pathways. A good example for a very homogeneous pathway is PPP as all its k_{app} values increased equally (all green within the pathway in Fig. 5) with faster growth, showing tightly controlled regulation over the whole pathway. However, glycolysis seems to have three groups of enzymes with different regulation principles: two groups of enzymes with similar (red) or faster (green) increase of k_{app} values compared to ribosomal proteins and one for which k_{app} showed no change at all. This confirms that control of glycolysis is realized through different levels and distributed over several parts to successfully balance the needs for substrate uptake, biomass monomer synthesis and energy generation, as generally believed. Another example of divergent protein regulation within a pathway is that of purine synthesis which possesses two large groups (red and green) and two small sets of genes. Currently, we have no possible explanation for the presence of these groups. This will be analyzed further in detail since it might be an important observation as the need for NADPH through purine synthesis pathways increases with rising μ[11] and hence, the question arises: why does the PPP pathway facilitate unequal regulation?

Efficiency of energy production pathways (E_{ATP})

Increasing transcription and translation rates with rising k_{app} of metabolic enzymes at faster growth contribute to the increase of metabolic efficiency (see above). To gain further insights into regulation of energy metabolism, we calculated the efficiency of energy production pathways (E_{ATP})—ATP produced in the pathway per ATP spent for synthesis of the pathway proteins (molecules-ATP/molecules-ATP)—for the main ATP yielding pathways under aerobic growth of *E. coli*: glycolysis, acetate synthesis and the TCA cycle coupled to the respiratory chain (RC) (Table 1). We believe that the cost of protein synthesis of the ATP producing pathways might be a relevant factor in optimization of the overall strategies of energy production since >45% of total ATP for cell proliferation is used for protein synthesis.[55] The most efficient energy generating pathway over the studied range of μ is the TCA cycle + RC, for which E_{ATP} varies between 105–152 molecules-ATP/molecules-ATP and is ~5- to 7-fold higher compared to glycolysis (Table 1). In addition to these pathways, *E. coli* starts to generate additional ATP through acetate excretion after disruption of the PTA-ACS cycle at μ = 0.27 h^{-1}.[13] Interestingly, carbon wasting into acetate seems to be beneficial for *E. coli* since E_{ATP} from acetate synthesis exceeds glycolysis at μ = 0.4 h^{-1} and even surpasses TCA cycle + RC near μ_{max} (141 *vs.* 132 molecules-ATP/molecules-ATP). The abrupt increase in E_{ATP} of acetate synthesis between μ = 0.4 and 0.48 h^{-1} is possibly coupled to the similar fast increase of the pathway's average k_{app} (Fig. 3). This observation provides a new angle for the potential rationale of acetate excretion for ATP generation[56,57] at the level of the whole cell and is a good example of how new biological knowledge can emerge from integration of genome-wide multi-level quantitative data.

Analysis of COG functional classes, pathways, individual genes and pm characteristics

We further analyzed our omics data in more detail by grouping all the quantified mRNAs, proteins and synthesis costs for each protein (*n*ATP, see Experimental) into COG functional classes.[36] Although, translational machinery proteins (J) were by far the most abundant COG group (23–36% of total proteome; $p < 0.0001$), synthesis cost of group J proteins was exceeded at all μ by the cost of proteins related to energy generation (C), carbohydrate (G) and amino acid (E) metabolism (Fig. S8, ESI†; $p < 0.0011$). This is expected since increased substrate utilization, amino acid synthesis and energy generation are needed for faster growth. The distribution of mRNA abundances was similar to proteins.

Turning attention to pathways, proteins of glycolysis, the TCA cycle, GS and purine synthesis accounted for both the highest concentration (~19%) and synthesis cost (~27%) of the total proteome (Fig. S9, ESI†). In agreement with flux measurements[8,10,11] and rising demands for energy generation and nucleotide synthesis for faster growth, protein expression in GS, gluconeogenesis and the TCA cycle, purine synthesis was down- and up-regulated, respectively, with rising μ.

Ribosomes, mainly composed of ribosomal proteins and rRNA, are one of the most important molecules enabling the cells to grow bigger and faster. It has been stated that ribosomal protein levels are insensitive to growth rate.[53] However, our data demonstrate 3.4-fold higher median ribosomal protein

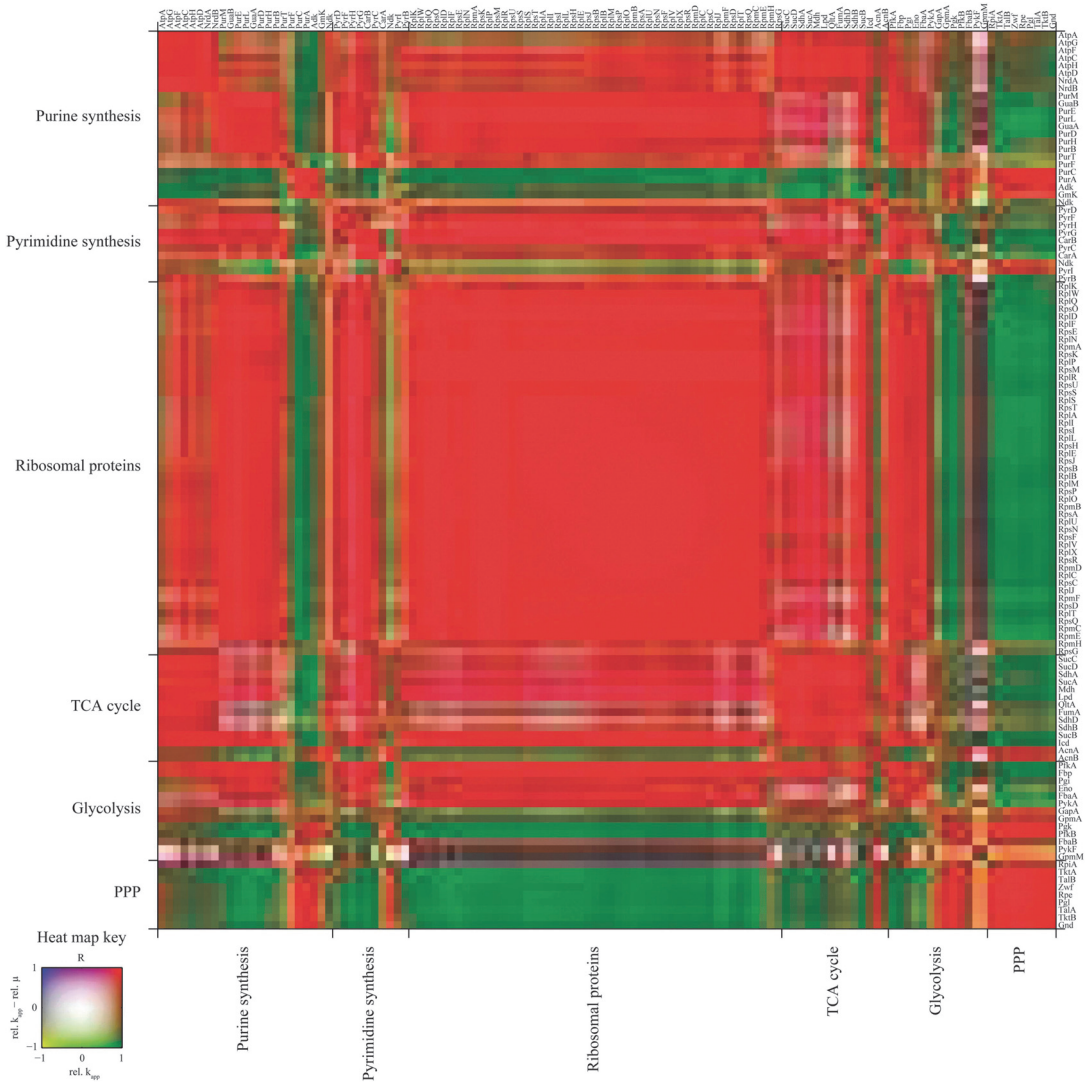

Fig. 5 Co-regulation analysis of protein synthesis and selected metabolic and biosynthetic pathways. Proteins belonging to respective pathways can be seen to the right and above the heat map. See Table S7 (ESI†) for the genes assigned to pathways according to the EcoCyc database.[46] Refer to the heat map key for interpretation of heat map colors: relative change of each protein's k_{app} with rising μ (rel. k_{app}, k_{app} at each μ compared to the value at reference $\mu = 0.11\ h^{-1}$) is compared to derive a correlation (R), indicated by color, between two proteins (horizontal comparison in heat map key); relative change of each protein's k_{app} with rising μ minus relative increase of μ (rel. k_{app} − rel. μ) is compared to derive a correlation (R), indicated by color, between two proteins (vertical comparison in heat map key). A square in the heat map denotes whether the two proteins' k_{app} change in the same direction with rising μ (red to green) or not (blue to yellow) and whether changes of the two proteins' relative k_{app} values are higher than relative increase of μ (blue to red) or not (yellow to green). For example, a red square denotes that the two proteins' k_{app} increase with rising μ and this change is for both proteins higher or lower than the increase in μ. R, PCC.

abundances (7164 *vs.* 24 509 molecules per cell) in faster growing cells while accounting for 9–16% of the total protein mass which is close to the estimated value of 21% previously reported for faster growing log-phase *E. coli* cells.[6]

We applied covariance analysis also to calculate and compare *n*ATP of different proteins for *E. coli* to achieve faster growth (Fig. 6). The most costly protein for *E. coli* by far with increasing μ was MetE, probably expressed to counteract the detrimental accumulation of homocysteine[58] due to increasing acid stress after acetate overflow switch.[13] MetE was followed mainly by ribosomal and amino acid metabolism-related proteins. On the other hand, *E. coli* 'saved' the most energy by

Table 1 Efficiency of energy production pathways (E_{ATP}) for the main ATP generating pathways under aerobic growth of *E. coli*

Pathway	$\mu = 0.11\ h^{-1}$	$\mu = 0.20\ h^{-1}$	$\mu = 0.30\ h^{-1}$	$\mu = 0.40\ h^{-1}$	$\mu = 0.49\ h^{-1}$
Glycolysis	22.6	21.2	20.3	18.8	19.7
TCA cycle + RC	151.1	116.9	105.5	105.5	132
Acetate synthesis	ND	ND	7.3	41.9	141.2

E_{ATP} is calculated as ATP produced in the pathway per ATP spent for synthesis of the pathway proteins (molecules–ATP/molecules–ATP). ATP produced in the pathway was calculated based on all the ATP producing and consuming fluxes in the respective pathway previously determined by metabolic flux analysis.[11] ATP spent for synthesis of the pathway proteins was calculated by summing all the quantified protein concentrations in the pathway, assuming an average protein length of 300 amino acids and 4.306 ATP for the cost of polymerization of one amino acid by the ribosome.[54] See Table S7 (ESI) for the genes assigned to pathways according to the EcoCyc database.[46] RC, respiratory chain. ND, not determined since the calculation would be inaccurate due to lack of exact data for both Pta-AckA and Acs fluxes in the PTA-ACS cycle before the start of overflow metabolism of acetate at μ = 0.27 h^{-1},[13] while functioning only as an intracellular futile cycle.

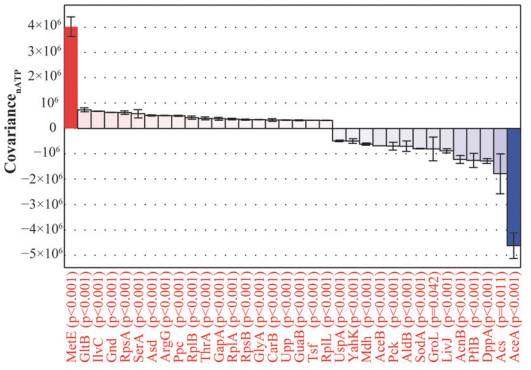

Fig. 6 Covariance analysis of protein synthesis cost (*n*ATP) for the 20 most and least expensive proteins for *E. coli* to achieve faster growth. Covariance values are calculated between each protein's *n*ATP values and μ. Red and blue colored bars indicate increasing and decreasing *n*ATP, respectively, with rising μ. Change of *n*ATP is statistically significant for all shown proteins since their covariance values are different from zero. Refer to Fig. 2A legend for description of error bars and statistical analysis.

repression of GS enzymes AceA and AceB, acetate scavenging Acs and enzymes involved in utilization of alternative substrates, and several stress response proteins. This is in line with the activation of carbon catabolite repression mediated decrease of flux through GS and disruption of acetate cycling with increasing μ.[8,10,13] Data for all genes are provided in Table S2 (ESI†).

Non-constant transcription rate and variable patterns of pm shown above can be influenced by half-life of proteins.[34] Stability of proteins is affected by particular amino acids present at the N-terminal end of proteins, known as the N-end rule. This has also been reported for *E. coli* with the destabilizing amino acids being Arg, Lys, Leu, Phe, Tyr and Trp.[59] Our analysis of whole protein sequences throughout the range of μ confirms the latter as Arg, Leu, Trp and Glu showed statistically significant enrichment ($p < 0.05$) in proteins both with the lowest pm ratios and concentration (Fig. S10, ESI†), demonstrating the influence of protein degradation on pm levels.

As cells have evolved in energy-limited conditions, the amino acid composition of the proteome should be biased towards containing more amino acids with lower synthesis costs. Indeed, highly expressed proteins in *E. coli* and *B. subtilis* contain more 'cheap' (Glu, Asp, Gly) and less 'costly' (Trp, Phe, His, Cys, Leu) amino acids.[60] Our proteome-wide data show the same as the most and least abundant proteins contained more 'cheap' and more 'costly' amino acids, respectively (Fig. S11, ESI†). This bias is also seen in the amino acid distribution of *E. coli* biomass which includes more 'cheap' amino acids.[11]

Codon bias (*i.e.* non-random occurrence of codons for coding amino acids) is believed to be the main mechanism for the cell to maximize translation efficiency[61] and lately genome-wide correlation between codon adaptation index (CAI),[62] a common estimate for codon bias, and protein expression levels has been presented for *E. coli*.[37,63–67] However, experiments using synthetic genes show that neither local nor global codon bias have significant effects on mRNA or protein levels,[68] although these results have later been objected.[69] Our results are in accord with the previous genome-wide studies as CAI values correlated with protein concentration and pm (Fig. S12, ESI†), and interestingly increased slightly with rising μ alluding to possible growing pressure for higher translation efficiency for achieving faster growth.

Analysis of mRNA and protein dynamics organized in transcription units and protein complexes

As both transcriptome and proteome data were quantified, we analyzed expression levels of mRNAs and proteins organized in transcription units (TUs) and protein complexes.[46] Firstly, we looked if mRNAs and proteins associated with TUs have less variance within their abundances compared to the whole data set at all μ. CV among all quantified mRNAs and proteins was around 4- and 3-fold higher than the CV within TUs, respectively (Fig. S13, ESI†) like in *Leptospira interrogans*.[70] Protein concentrations within protein complexes are similarly 3-fold less noisy compared to the whole proteome (Fig. S14, ESI†), also seen in yeast.[71] The same phenomena for both cases can also be seen when comparing the differences in expression change (Fig. S13 and S14, ESI†).

Next, we asked if mRNA and protein abundances associated with *E. coli* TUs show a similar kind of 'staircase-behavior' of higher transcript expression at the 5′ end of operons seen in half of the 139 polycistronic operons in *M. pneumoniae*[72] and globally in *Streptomyces coelicolor*.[73] At the protein level, staircase-like expression exists only for ~5% of *L. interrogans* operons.[70] We detected staircase-like expression for 28 and 51% of polycistronic TUs on average over the studied range of μ at mRNA and protein levels, respectively (Fig. S15, ESI†). Dividing TUs according to their staircase-behavior type (see Experimental, and Fig. S16, ESI†) showed close proportions between groups 'down', 'up' and 'others' (Fig. S15, ESI†) whereas no 'up'-like expression

was specifically detected for *M. pneumoniae* operons.[72] Interestingly, *E. coli* does not seem to compensate the observed 'operon polarity' of transcripts on the protein level (Fig. S17, ESI†) as seen in *M. pneumoniae*.[40] These results further demonstrate the additional regulatory flexibility of transcription in bacteria similar to eukaryotes. Further evidence for these gene expression regulation mechanisms being conserved in bacteria came from the observation that mRNA and protein concentrations correlated better in longer TUs and for genes located at the 5′ end (Fig. S18, ESI†), similar to *M. pneumoniae*.[40]

Experimental

Bacterial strain, growth medium and continuous cultivation conditions

E. coli K-12 MG1655 (λ- F- *rph-1Fnr+*; Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), DSM No.18039) strain was grown on a defined minimal medium[9] with 4.5 g L^{-1} α-D-glucose. Three independent A-stat[23] continuous cultivation experiments were performed under the following conditions: temperature 37 °C, pH 7, agitation speed of 800 rpm, and aerobic conditions (air flow rate 150 ml min^{-1}). The A-stat algorithm used was: $D = D_0 + a_D \times t$, where D_0 is the initial dilution rate of chemostat (0.1 h^{-1}), a_D is the acceleration rate in the A-stat phase (0.01 h^{-2}), and t is the time from the start of acceleration (h). A detailed description of cultivation conditions and growth characteristics in these experiments has been reported previously.[13]

Experimental data

Specific flux rates (mmol $gDCW^{-1}$ h^{-1}) determined by metabolomics and metabolic flux analysis in the same cultivation experiments were taken from Valgepea *et al.*[11] All flux values used in calculations except fluxes directly depending on biomass monomer composition are average of three A-stats.

Genome-wide transcriptome analysis at μ = 0.11; 0.21; 0.31; 0.40; 0.48 h^{-1} was performed previously using Agilent DNA microarrays (GEO reference series: GSE23920) for around 4300 mRNAs.[13] Transcriptome analysis was conducted in one A-stat with six technical replicates for reference sample at μ = 0.11 h^{-1}. Oligo spot intensities of the Agilent platform can be used as a proxy for mRNA abundances, since spot intensities and mRNA abundances correlate perfectly (R^2 = 1.00; see Fig. 6 in Agilent Application Note 5989-9159EN). Therefore, for each microarray, average spot intensities of three different mRNA oligos were summed excluding ribosomal and transport RNA oligos corresponding to the total amount of mRNA. To convert this value from the spot intensity unit into g total mRNA per g of dry cell weight (DCW), total RNA % in DCW was determined (Table S1, ESI†) since it is difficult to accurately determine mRNA % in DCW. Recently, mRNA % from total RNA was shown to be 4.9% in late-exponential *E. coli* BW25113 cells[74] and constant in exponential, stationary phase and heat shock-treated *E. coli* K-12 MG1655 cells using RNA-seq in both cases.[75] In addition, we have measured a 5% yield of cDNA synthesis with poly-A primers from total RNA in *S. cerevisiae* continuous cultures at various μ. Therefore, we assumed the best estimate for mRNA content in total RNA to be 5% under the studied range of μ. mRNA molecule numbers in gDCW were calculated by the fraction of each mRNA's spot intensity from the summed spot intensities (described above) corresponding to the estimated total amount of mRNA in gDCW taking into account the molecular weight of each mRNA. Finally, mRNA abundances (molecules per cell) were calculated from the previous values using biomass concentration and cell counts in the culture broth. Biomass concentration expressed as gDCW L^{-1} was determined gravimetrically as described by Nahku *et al.*[9] Cell counts were measured by incubating five replicate LB-agar plates at 37 °C for ~11–12 h and expressed as CFU ml^{-1}, equivalent to cell ml^{-1}. Data can be seen in Table S1 (ESI†). Only mRNAs which spot intensities had a signal-to-noise ratio higher than three (spot intensity > 100) were used in covariance and staircase-behavior type analyses to exclude genes with high fold changes between different μ caused by very low spot intensities.

Absolute proteome quantification was performed for 1185 proteins at μ = 0.11 h^{-1} in two A-stats using the iBAQ approach[24] and published in Arike *et al.*[25] Protein abundances (molecules per cell) for μ = 0.20; 0.30; 0.40; 0.49 h^{-1} were calculated based on relative protein expression data published previously in Valgepea *et al.*[13] covering a μ range of 0.11–0.49 h^{-1} taking into account μ-dependent cell counts and total protein % in DCW (Table S1, ESI†) measured by the Lowry method.[76] All protein abundances and concentrations used in calculations are mean of two independent A-stats.

mRNA and protein abundances (molecules per cell) were converted into intracellular concentrations (molecules per fL of biomass) using cell volume at all μ estimated from biomass concentration and cell counts (Table S1, ESI†) as follows:

$$\text{intracellular concentration}\left(\frac{\text{molecules}}{\text{fL}}\right) = \text{abundance}\left(\frac{\text{molecules}}{\text{cell}}\right) \times \left(\frac{1}{\text{cell volume (fL)}}\right) \quad (1)$$

Calculations

Protein-per-mRNA (pm) ratios were calculated for all genes with quantified mRNA and protein concentrations as follows:

$$\text{pm} = \frac{\text{protein}_i}{\text{mRNA}_i}, \quad (2)$$

where protein_i and mRNA_i are individual mRNA and protein concentrations (molecules per fL), respectively.

Apparent catalytic rates (k_{app}, s^{-1}) were calculated for 191 metabolic enzymes associated with catalyzing the fluxes calculated previously by metabolic flux analysis[11] and 52 ribosomal proteins assuming each protein chain being catalytically active. Proteins were assigned to fluxes according to the EcoCyc database.[46] The sum of all amino acid synthesis fluxes was used as the flux catalyzed by ribosomal proteins. k_{app} values were calculated as follows:

$$k_{\text{app}_i} = \frac{\text{specific flux rate}_i}{\text{protein}_i} \times \frac{N_A \times 0.3}{10^{12}}, \quad (3)$$

where specific flux rate$_i$ (mol gDCW^{-1} s^{-1}) is the flux catalyzed by protein$_i$ (molecules per fL of biomass), N_A is the Avogadro number ($\sim 6.02 \times 10^{23}$), 0.3 is the fraction of dry cell mass in one gram of biomass and 10^{12} is the conversion factor from fL to g assuming a buoyant density of 1 g ml^{-1} for an *E. coli* cell.

RNAP synthesis rate (nucleotides per RNAP molecules per s) at each μ was calculated as follows:

$$\text{RNAP synthesis rate} = \frac{\text{total mRNA length} \times \mu}{\text{RNAP}} \times \frac{0.3}{10^{12}}, \quad (4)$$

where total mRNA length (nucleotides per gDCW) is the sum of all quantified and hence transcribed mRNA lengths, μ (s^{-1}) is specific growth rate, RNAP (molecules per fL of biomass) is the average concentration of the two RNAP β subunit proteins RpoB and RpoC, 0.3 is the fraction of dry cell mass in one gram of biomass and 10^{12} is the conversion factor from fL to g assuming a buoyant density of 1 g ml^{-1} for an *E. coli* cell.

Covariance analysis

1. Covariance. We applied covariance as a measure of how gene expression levels (mRNA and protein concentrations) change with rising μ since covariance describes both the direction and magnitude of the changes, making it a more suitable statistic in this work since absolute quantitative data were acquired. Firstly, we used covariance to estimate mRNA and protein expression burden with increasing μ for *E. coli*. Secondly, covariance was used to calculate protein synthesis costs (*n*ATP) over the whole range of μ. Lastly, covariance analysis was applied to determine the regulation levels of gene expression using mRNA, protein and specific flux rates (*e.g.* TR, post-TR, TL, post-TL) as described below.

Covariance (COV) was calculated according to the formula:

$$\text{COV} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \quad (5)$$

for mRNA expression burden as:

$$\text{COV}_{\text{mRNA}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\text{mRNA}_i - \overline{\text{mRNA}}\right)(\mu_i - \bar{\mu}) \quad (6)$$

for protein expression burden as:

$$\text{COV}_{\text{protein}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\text{prot}_i - \overline{\text{prot}}\right)(\mu_i - \bar{\mu}) \quad (7)$$

for *n*ATP as:

$$\text{COV}_{\text{nATP}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(n\text{ATP}_i - \overline{n\text{ATP}}\right)(\mu_i - \bar{\mu}) \quad (8)$$

for gene expression regulation analysis at protein and mRNA level (pm) as:

$$\text{COV}_{\text{pm}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\text{prot}_i/\text{mRNA}_i - \overline{\text{prot}/\text{mRNA}}\right)(\mu_i - \bar{\mu}) \quad (9)$$

for gene expression regulation analysis at specific flux rate and protein level (k_{app}) as:

$$\text{COV}_{k_{\text{app}}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\text{spec flux}_i/\text{prot}_i - \overline{\text{spec flux}/\text{prot}}\right)(\mu_i - \bar{\mu}), \quad (10)$$

where mRNA, prot and spec flux represent mRNA and protein concentrations, and specific flux rates measured at their respective μ. See below for *n*ATP calculation.

2. Uncertainty of covariance. Covariance analysis requires uncertainty values of covariance for the statistical hypothesis testing of covariance values being statistically different from zero or being statistically zero. Uncertainty of covariance was calculated according to the covariance formula shown above.

The relative uncertainty of covariance can be found by:

$$\begin{aligned}\frac{u_{\text{COV}}}{\text{COV}} &= \sqrt{\sum_{i=1}^{n}\left(\frac{u_{(x_i-\bar{x})}}{(x_i-\bar{x})}\right)^2 + \sum_{i=1}^{n}\left(\frac{u_{(y_i-\bar{y})}}{(y_i-\bar{y})}\right)^2} \\ &= \sqrt{\sum_{i=1}^{n}\left(\frac{u_{x_i}}{x_i}\right)^2 + \left(\frac{u_{\bar{x}}}{\bar{x}}\right)^2 + \sum_{i=1}^{n}\left(\frac{u_{y_i}}{y_i}\right)^2 + \left(\frac{u_{\bar{y}}}{\bar{y}}\right)^2},\end{aligned} \quad (11)$$

where u is the absolute uncertainty.

As mean values of $\bar{x}$ and $\bar{y}$ are calculated according to formulas:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i \quad (12)$$

and

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n}y_i, \quad (13)$$

then the uncertainty of mean value can be found by:

$$\frac{u_{\bar{x}}}{\bar{x}} = \sqrt{\sum_{i=1}^{n}\left(\frac{u_{x_i}}{x_i}\right)^2} \quad (14)$$

and

$$\frac{u_{\bar{y}}}{\bar{y}} = \sqrt{\sum_{i=1}^{n}\left(\frac{u_{y_i}}{y_i}\right)^2} \quad (15)$$

Therefore, the relative uncertainty of covariance is equal to:

$$\frac{u_{\text{COV}}}{\text{COV}} = \sqrt{2\sum_{i=1}^{n}\left(\frac{u_{x_i}}{x_i}\right)^2 + 2\sum_{i=1}^{n}\left(\frac{u_{y_i}}{y_i}\right)^2} \quad (16)$$

The uncertainty of μ is considered to be zero since all the compared data (mRNA and protein concentration, specific flux rate) were acquired at the same μ. Therefore, the member with y in the second half of the equation becomes zero.

As pm and k_{app} are ratios between protein and mRNA, and specific flux and protein, the uncertainty of pm or k_{app} is made of uncertainty estimations of the members of the particular ratio. Uncertainty of mRNA concentrations was expressed

through 95% CI calculated for each mRNA based on six technical replicates at $\mu = 0.11\ h^{-1}$. Uncertainty of protein concentrations was expressed through average absolute deviation calculated for each protein based on proteome quantification in two independent cultivation experiments. As the uncertainty of the specific flux rate was not available for all the genes (biomass monomer composition was not determined in all independent A-stat experiments) it was not taken into account in the uncertainty value of k_{app}, as otherwise it would have given unequal statistical preference to the genes in which uncertainty of the specific flux rate was not estimated.

Therefore, uncertainty estimations of pm and k_{app} were the following:

$$\frac{u_{\mathrm{COV_{pm}}}}{\mathrm{COV_{pm}}} = \sqrt{2\sum_{i=1}^{n}\left(\frac{u_{x_{\mathrm{prot},i}}}{x_{\mathrm{prot},i}}\right)^2 + 2\sum_{i=1}^{n}\left(\frac{u_{x_{\mathrm{mRNA},i}}}{x_{\mathrm{mRNA},i}}\right)^2} \quad (17)$$

$$\frac{u_{\mathrm{COV}_{k_{app}}}}{\mathrm{COV}_{k_{app}}} = \sqrt{2\sum_{i=1}^{n}\left(\frac{u_{x_{\mathrm{prot},i}}}{x_{\mathrm{prot},i}}\right)^2} \quad (18)$$

*n*ATP was calculated according to the formula:

$$x_{n\mathrm{ATP}} = x_{\mathrm{prot}} \times (x_{n\mathrm{AA}} - 1) \times 4.306, \quad (19)$$

where x_{prot} is the protein concentration measured at the respective μ, $x_{n\mathrm{AA}}$ is the number of amino acids in the protein and 4.306 represents the cost of polymerization of one amino acid into the growing peptide chain by the ribosome in ATP.[54] As the length of the protein is considered a constant value, the uncertainty of *n*ATP ($x_{n\mathrm{ATP}}$) is equal to relative uncertainty of protein expression:

$$\frac{u_{x_{n\mathrm{ATP}}}}{x_{n\mathrm{ATP}}} = \frac{u_{x_{\mathrm{prot}}}}{x_{\mathrm{prot}}} \quad (20)$$

Uncertainty of *n*ATP covariance is the following:

$$\frac{u_{\mathrm{COV}_{n\mathrm{ATP}}}}{\mathrm{COV}_{n\mathrm{ATP}}} = \sqrt{2\sum_{i=1}^{n}\left(\frac{u_{x_{n\mathrm{ATP},i}}}{x_{n\mathrm{ATP},i}}\right)^2} = \sqrt{2\sum_{i=1}^{n}\left(\frac{u_{x_{\mathrm{prot},i}}}{x_{\mathrm{prot},i}}\right)^2} \quad (21)$$

3. **Test of significance.** Calculated covariance values were subjected to statistical hypothesis testing. Firstly, the hypothesis that absolute values of covariance are different from zero at a statistically significant level was tested. One sided *Z*-test was applied to compare absolute values of covariance to zero. Uncertainty estimation calculated as described above was used as a nuisance parameter. Genes with significance levels below 0.05 were considered as possibly statistically significant from zero and subjected to false discovery rate (FDR) filtering at level $\alpha = 0.05$ according to the Benjamini–Hochberg method.[77] Genes with significance levels below the threshold value of FDR filtering were considered statistically significantly different from zero. The rest of the covariance values do not differ significantly from zero. This group of genes is made of genes either whose covariance value is zero, or have very high uncertainty values. To determine which covariance values are actually equal to zero, additional hypothesis testing was applied. Firstly, it was expected that in most of the cases when a covariance value does not differ from zero at a statistically significant level, it is because of the fact that the covariance value is actually zero and not because the uncertainty value of covariance is very high. Probability density of absolute values of covariance is expected to follow one sided normal distribution. Covariance values below 95% CI of this distribution are considered zeros. The *Z*-test was applied to test the hypothesis that absolute values of covariance are below the confidence interval. Results of the *Z*-test were again subjected to FDR filtering. The covariance value was considered to be zero if the value was below the threshold limit of FDR filtering.

Eventually genes were divided into three groups. One group corresponded to genes with a covariance value statistically higher than zero. In gene expression regulation analysis, these genes' expression regulation level is referred to as post-transcriptional or translational for protein and mRNA, and post-translational for specific flux rate and protein regulation levels. The second group corresponded to genes with a covariance value equal to zero at the statistically significant level. Again in gene expression regulation analysis, these genes' expression regulation level is referred to as transcriptional for protein and mRNA, and translational for specific flux rates and protein regulation levels. The rest of the genes were described by such a high uncertainty level of covariance that it was impossible to determine their nature towards zero. Identical hypothesis testing was applied to covariance values of pm and k_{app}.

Analysis of staircase-behavior type gene expression in TUs

Analysis of staircase-behavior type gene expression in TUs was analyzed only for polycistronic TUs. All quantified mRNAs and proteins were divided into TUs according to the EcoCyc database[46] using a script developed in-house. TU mRNA and protein expression levels were divided into "no staircase" and staircase-like expression behavior types "up", "down" and "others" (see Fig. S16, ESI† for a visual description). A TU was classified as "no staircase" if at least half of its consecutive genes did not show statistically significant expression difference at significance level $p < 0.05$ in a *Z*-test. Uncertainties of mRNA and protein concentrations used in the *Z*-test as nuisance parameters are described in the Covariance analysis section above. All the remaining TUs were considered to show a staircase-like expression. Staircase-behavior type was classified as "up" or "down" if at least half of its consecutive genes were differentially expressed at higher or lower levels, respectively, in the mRNA emerging direction in transcription (5′ → 3′). The remaining TUs were classified as showing an "others" staircase-behavior type.

Conclusion

This first μ-dependent absolute quantitative *E. coli* multi-omics data set generated new knowledge about genome-wide regulation levels of gene expression for protein and flux control, their mRNA and protein synthesis burden, and various other molecular,

energetic and metabolic aspects of the gene expression cascade. More importantly, we showed that cells achieve faster growth predominantly by increasing catalytic and translation rates of proteins which supports the recent conclusion that transcriptional control in metabolism might not be as dominant as once thought.[18,43,48] These observations together with the observed increase in efficiency of energy generation most probably contribute to the ability of *E. coli* to maintain constant biomass yield under increased carbon wasting with rising μ through increased metabolic efficiency.[11] We believe that these findings together with our multi-level quantitative data set determined in defined physiological states of *E. coli* can advance modeling approaches (*e.g.* addition of non-constant transcription and translation rates into growth-rate coupled models such as those presented in Scott *et al.*[30] and Lerman *et al.*[32]) and add valuable information to the much needed better description and understanding of growth rate regulation through coordinated operational processes at the whole cell level[16] and quantitative understanding of biological systems overall.[19] Conclusively, the observations reported in this study could also lead to more efficient metabolic engineering of industrial strains since new knowledge regarding protein and flux control levels was presented.

Acknowledgements

The authors thank Ranno Nahku for help with absolute transcriptome quantification, Klim Evdokimov with data handling, and Petri-Jaan Lahtvee and Karl Peebo for critical reading of the manuscript. The financial support for this work was provided by the European Regional Development Fund project EU29994; Ministry of Education, Estonia, through the grant SF0140090s08 and Estonian Science Foundation through grants G8165 and G9192.

References

1 F. Jacob and J. Monod, *J. Mol. Biol.*, 1961, **3**, 318–356.
2 U. Sauer, *Mol. Syst. Biol.*, 2006, **2**, 62.
3 B. H. ter Kuile and H. V. Westerhoff, *FEBS Lett.*, 2001, **500**, 169–171.
4 K. van Eunen, S. Rossell, J. Bouwman, H. V. Westerhoff and B. M. Bakker, *Methods Enzymol.*, 2011, **500**, 571–595.
5 P. Daran-Lapujade, S. Rossell, W. M. van Gulik, M. A. H. Luttik, M. J. L de Groot, M. Slijper, A. J. R. Heck, J.-M. Daran, J. H. de Winde, H. V. Westerhoff, J. T. Pronk and B. M. Bakker, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 15753–15758.
6 H. Bremer and P. Dennis, in *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology, ed. F. Neidhardt, R. Curtiss III, J. Ingraham, E. Lin, K. Low, B. Magasanik, W. Reznikoff, M. Riley, M. Schaechter and H. Umbarger, ASM Press, Washington, DC, 2nd edn, 1996, pp. 1553–1569.
7 G. N. Vemuri, E. Altman, D. P. Sangurdekar, A. B. Khodursky and M. A. Eiteman, *Appl. Environ. Microbiol.*, 2006, **72**, 3653–3661.
8 N. Ishii, K. Nakahigashi, T. Baba, M. Robert, T. Soga, A. Kanai, T. Hirasawa, M. Naba, K. Hirai, A. Hoque, P. Y. Ho, Y. Kakazu, K. Sugawara, S. Igarashi, S. Harada, T. Masuda, N. Sugiyama, T. Togashi, M. Hasegawa, Y. Takai, K. Yugi, K. Arakawa, N. Iwata, Y. Toya, Y. Nakayama, T. Nishioka, K. Shimizu, H. Mori and M. Tomita, *Science*, 2007, **316**, 593–597.
9 R. Nahku, K. Valgepea, P.-J. Lahtvee, S. Erm, K. Abner, K. Adamberg and R. Vilu, *J. Biotechnol.*, 2010, **45**, 60–65.
10 A. Nanchen, A. Schicker and U. Sauer, *Appl. Environ. Microbiol.*, 2006, **72**, 1164–1172.
11 K. Valgepea, K. Adamberg and R. Vilu, *BMC Syst. Biol.*, 2011, **5**, 106.
12 P.-J. Lahtvee, K. Adamberg, L. Arike, R. Nahku, K. Aller and R. Vilu, *Microb. Cell Fact.*, 2011, **10**, 12.
13 K. Valgepea, K. Adamberg, R. Nahku, P.-J. Lahtvee, L. Arike and R. Vilu, *BMC Syst. Biol.*, 2010, **4**, 166.
14 J. A. Mongold and R. E. Lenski, *J. Bacteriol.*, 1996, **178**, 5333–5334.
15 B. Volkmer and M. Heinemann, *PLoS One*, 2011, **6**, e23126.
16 F. C. Neidhardt, *J. Bacteriol.*, 1999, **181**, 7405–7408.
17 U. Sauer, M. Heinemann and N. Zamboni, *Science*, 2007, **316**, 550–551.
18 M. Heinemann and U. Sauer, *Curr. Opin. Microbiol.*, 2010, **13**, 337–343.
19 R. Phillips and R. Milo, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 21465–21471.
20 M. Scott and T. Hwa, *Curr. Opin. Biotechnol.*, 2011, **22**, 559–565.
21 P. A. Hoskisson and G. Hobbs, *Microbiology*, 2005, **151**, 3153–3159.
22 A. T. Bull, *J. Ind. Microbiol. Biotechnol.*, 2010, **37**, 993–1021.
23 T. Paalme, A. Kahru, R. Elken, K. Vanatalu, K. Tiisma and R. Vilu, *J. Microbiol. Methods*, 1995, **24**, 145–153.
24 B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen and M. Selbach, *Nature*, 2011, **473**, 337–342.
25 L. Arike, K. Valgepea, L. Peil, R. Nahku, K. Adamberg and R. Vilu, *J. Proteomics*, 2012, **75**, 5437–5448.
26 A. Otto, J. Bernhardt, H. Meyer, M. Schaffer, F.-A. Herbst, J. Siebourg, U. Mäder, M. Lalk, M. Hecker and D. Becher, *Nat. Commun.*, 2010, **1**, 137.
27 C. Vogel and E. M. Marcotte, *Nat. Rev. Genet.*, 2012, **13**, 227–232.
28 S. Marguerat, A. Schmidt, S. Codlin, W. Chen, R. Aebersold and J. Bähler, *Cell*, 2012, **151**, 671–683.
29 M. Schaechter, O. Maaløe and N. O. Kjeldgaard, *J. Gen. Microbiol.*, 1958, **19**, 592–606.
30 M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang and T. Hwa, *Science*, 2010, **330**, 1099–1102.
31 S. Klumpp, Z. Zhang and T. Hwa, *Cell*, 2009, **139**, 1366–1375.
32 J. A. Lerman, D. R. Hyduke, H. Latif, V. A. Portnoy, N. E. Lewis, J. D. Orth, A. C. Schrimpe-Rutledge, R. D. Smith, J. N Adkins, K. Zengler and B. Ø. Palsson, *Nat. Commun.*, 2012, **3**, 929.
33 R. de Sousa Abreu, L. O. Penalva, E. M. Marcotte and C. Vogel, *Mol. BioSyst.*, 2009, **5**, 1512–1526.
34 T. Maier, M. Güell and L. Serrano, *FEBS Lett.*, 2009, **583**, 3966–3973.
35 W. Zhang, F. Li and L. Nie, *Microbiology*, 2010, **156**, 287–301.

36 R. L. Tatusov, N. D. Fedorova, J. D. Jackson, A. R. Jacobs, B. Kiryutin, E. V. Koonin, D. M. Krylov, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, S. Smirnov, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin and D. A. Natale, *BMC Bioinf.*, 2003, **4**, 41.

37 P. Lu, C. Vogel, R. Wang, X. Yao and E. M. Marcotte, *Nat. Biotechnol.*, 2007, **25**, 117–124.

38 Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili and X. S. Xie, *Science*, 2010, **329**, 533–538.

39 K. Adamberg, A. Seiman and R. Vilu, *PLoS One*, 2012, **7**, e48223.

40 T. Maier, A. Schmidt, M. Güell, S. Kühner, A.-C. Gavin, R. Aebersold and L. Serrano, *Mol. Syst. Biol.*, 2011, **7**, 511.

41 J. M. Buescher, W. Liebermeister, M. Jules, M. Uhr, J. Muntel, E. Botella, B. Hessling, R. J. Kleijn, L. Le Chat, F. Lecointe, U. Mader, P. Nicolas, S. Piersma, F. Rugheimer, D. Becher, P. Bessieres, E. Bidnenko, E. L. Denham, E. Dervyn, K. M. Devine, G. Doherty, S. Drulhe, L. Felicori, M. J. Fogg, A. Goelzer, A. Hansen, C. R. Harwood, M. Hecker, S. Hubner, C. Hultschig, H. Jarmer, E. Klipp, A. Leduc, P. Lewis, F. Molina, P. Noirot, S. Peres, N. Pigeonneau, S. Pohl, S. Rasmussen, B. Rinn, M. Schaffer, J. Schnidder, B. Schwikowski, J. M Van Dijl, P. Veiga, S. Walsh, A. J. Wilkinson, J. Stelling, S. Aymerich and U. Sauer, *Science*, 2012, **335**, 1099–1103.

42 A. Vazquez, Q. K. Beg, M. A. Demenezes, J. Ernst, Z. Bar-Joseph, A.-L. Barabási, L. G. Boros and Z. N Oltvai, *BMC Syst. Biol.*, 2008, **2**, 7.

43 L. Gerosa and U. Sauer, *Curr. Opin. Biotechnol.*, 2011, **22**, 566–575.

44 B. R. B. Haverkorn van Rijsewijk, A. Nanchen, S. Nallet, R. J. Kleijn and U. Sauer, *Mol. Syst. Biol.*, 2011, **7**, 477.

45 F. Wessely, M. Bartl, R. Guthke, P. Li, S. Schuster and C. Kaleta, *Mol. Syst. Biol.*, 2011, **7**, 515.

46 I. M. Keseler, A. Mackie, M. Peralta-Gil, A. Santos-Zavaleta, S. Gama-Castro, C. Bonavides-Martínez, C. Fulcher, A. M. Huerta, A. Kothari, M. Krummenacker, M. Latendresse, L. Muñiz-Rascado, Q. Ong, S. Paley, I. Schröder, A. G. Shearer, P. Subhraveti, M. Travers, D. Weerasinghe, V. Weiss, J. Collado-Viles, R. P. Gunsalus, I. Paulsen and P. D. Karp, *Nucleic Acids Res.*, 2013, **41**, D605–D612.

47 A. Lourenço, S. Carneiro, J. P. Pinto, M. Rocha, E. C. Ferreira and I. Rocha, *J. Integr. Bioinform.*, 2011, **8**, 183.

48 J. B. Plotkin, *Mol. Syst. Biol.*, 2010, **6**, 406.

49 A. Bar-Even, E. Noor and Y. Savir, *Biochemistry*, 2011, **50**, 4402–4410.

50 S. Sunya, F. Delvigne, J.-L. Uribelarrea, C. Molina-Jouve and N. Gorret, *Appl. Microbiol. Biotechnol.*, 2012, **95**, 1021–1034.

51 H. Taymaz-Nikerel, W. M. van Gulik and J. J. Heijnen, *Metab. Eng.*, 2011, **13**, 307–318.

52 R. Young and H. Bremer, *Biochem. J.*, 1976, **160**, 185–194.

53 D. J. Jin, C. Cagliero and Y. N. Zhou, *FEMS Microbiol. Rev.*, 2011, **36**, 269–287.

54 F. C. Neidhardt, J. Ingraham and M. Schaechter, *Physiology of the Bacterial Cell: A Molecular Approach*, Sinauer Associates, Sunderland, 1990, p. 96.

55 R. A. Cox, *Microbiology*, 2004, **150**, 1413–1426.

56 A. Kayser, J. Weber, V. Hecht and U. Rinas, *Microbiology*, 2005, **151**, 693–706.

57 A. J. Wolfe, *Microbiol. Mol. Biol. Rev.*, 2005, **69**, 12–50.

58 A. J. Roe, C. P. O'Byrne, D. McLaggan and I. R. Booth, *Microbiology*, 2002, **148**, 2215–2222.

59 J. W. Tobias, T. E. Shrader, G. Rocap and A. Varshavsky, *Science*, 1991, **254**, 1374–1377.

60 H. Akashi and T. Gojobori, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 3695–3700.

61 H. Gingold and Y. Pilpel, *Mol. Syst. Biol.*, 2011, **7**, 481.

62 P. M. Sharp and W. Li, *Nucleic Acids Res.*, 1987, **15**, 1281–1295.

63 B. Futcher, G. I. Latter, P. Monardo, C. S. Mclaughlin and I. Garrels, *Mol. Cell. Proteomics*, 1999, **19**, 7357–7368.

64 Y. Ishihama, T. Schmidt, J. Rappsilber, M. Mann, F. U. Hartl, M. J. Kerner and D. Frishman, *BMC Genomics*, 2008, **9**, 102.

65 G. Lithwick and H. Margalit, *Genome Res.*, 2003, **13**, 2665–2673.

66 T. Masuda, N. Saito, M. Tomita and Y. Ishihama, *Mol. Cell. Proteomics*, 2009, **8**, 2770–2777.

67 T. Tuller, Y. Y. Waldman, M. Kupiec and E. Ruppin, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 3645–3650.

68 G. Kudla, A. W. Murray, D. Tollervey and J. B. Plotkin, *Science*, 2009, **324**, 255–258.

69 F. Supek and T. Šmuc, *Genetics*, 2010, **185**, 1129–1134.

70 A. Schmidt, M. Beck, J. Malmström, H. Lam, M. Claassen, D. Campbell and R. Aebersold, *Mol. Syst. Biol.*, 2011, **7**, 510.

71 H. B. Fraser, A. E. Hirsh, G. Giaever, J. Kumm and M. B. Eisen, *PLoS Biol.*, 2004, **2**, e137.

72 M. Güell, V. van Noort, E. Yus, W.-H. Chen, J. Leigh-Bell, K. Michalodimitrakis, T. Yamada, M. Arumugam, T. Doerks, S. Kühner, M. Rode, M. Suyama, S. Schmidt, A.-C. Gavin, P. Bork and L. Serrano, *Science*, 2009, **326**, 1268–1271.

73 E. Laing, V. Mersinias, C. P. Smith and S. J. Hubbard, *Genome Biol.*, 2006, **7**, R46.

74 A. Shinhara, M. Matsui, K. Hiraoka, W. Nomura, R. Hirano, K. Nakahigashi, M. Tomita, H. Mori and A. Kanai, *BMC Genomics*, 2011, **12**, 428.

75 H. Yi, Y.-J. Cho, S. Won, J.-E. Lee, H. Jin Yu, S. Kim, G. P. Schroth, S. Luo and J. Chun, *Nucleic Acids Res.*, 2011, **39**, e140.

76 O. Lowry, N. Rosebrough, A. Lewis Farr and R. Randall, *J. Biol. Chem.*, 1951, **193**, 265–275.

77 Y. Benjamini and Y. Hochberg, *J. R. Statist. Soc. B*, 1995, **57**, 289–300.

CURRICULUM VITAE

Personal data

Name:	Kaspar Valgepea
Data and place of birth:	January 15, 1984 in Pärnu, Estonia
Citizenship:	Estonian
E-mail:	kaspar@tftak.eu

Education

2010 – present	Tallinn University of Technology, Tallinn, Estonia; Ph.D. student, Chemistry and Gene Technology
2006 – 2010	Tallinn University of Technology; M.Sc. *cum laude*, Applied Chemistry and Biotechnology
2003 – 2006	Tallinn University of Technology; B.Sc., Applied Chemistry and Biotechnology

Professional employment

2005 – present	Competence Center of Food and Fermentation Technologies, Tallinn, Estonia; Researcher
2012 – 2013	Genomatica, Inc, San Diego, USA; Professional Internship as a Visiting Scholar
2007 – 2008	Tokyo University of Electro-Communications, Tokyo, Japan; M.Sc. research

Special courses

2013	"Advanced Course on Metabolic Engineering and Systems Biology"; Chalmers University of Technology, Sweden
2012	"Advanced Course on Microbial Physiology and Fermentation Technology"; Delft University of Technology, The Netherlands
2011	"FEBS-SystemsX Advanced Lecture Course on Systems Biology"; Innsbruck, Austria

2009	"Flow Cytometry course"; University of York, United Kingdom

Teaching and supervising

2014	Supervisor for M.Sc. thesis of Karl Peebo "Specific growth rate dependent proteome dynamics of *Escherichia coli* in glucose minimal and defined rich medium"
2006 – present	Mentor of undergraduate and graduate students in the group of Prof. Raivo Vilu at the Competence Center of Food and Fermentation Technologies
2010	Lecturer and practical course advisor of "Advanced Microbial Fermentation Technologies", Tallinn University of Technology

Honors

2012	Baltic-American Freedom Foundation Professional Internship Program Scholarship
2011	Artur Lind Scholarship, Estonian Genome Centre Foundation
2010	2nd prize in Bio- and Environmental Sciences, Estonian National Students Contest of Research, M.Sc. level
2010	3rd prize in Life Sciences, Tallinn University of Technology Students Contest of Research, M.Sc. level
2009	Jaan Poska Scholarship
2007	Kristjan Jaak Scholarship

Language skills

Fluent	Estonian, English
Average	Japanese, Finnish
Basic	Russian

ELULOOKIRJELDUS

Isikuandmed

Nimi:	Kaspar Valgepea
Sünniaeg ja -koht:	15 Jaanuar 1984; Pärnu, Eesti
Kodakondsus:	eestlane
E-post:	kaspar@tftak.eu

Hariduskäik

2010 – hetkeni	Tallinna Tehnikaülikool, Tallinn, Eesti; Doktorant, Keemia ja geenitehnoloogia
2006 – 2010	Tallinna Tehnikaülikool; M.Sc. *cum laude*, Rakenduskeemia ja biotehnoloogia
2003 – 2006	Tallinna Tehnikaülikool; B.Sc., Rakenduskeemia ja biotehnoloogia

Teenistuskäik

2005 – hetkeni	Toidu- ja Fermentatsioonitehnoloogia Arenduskeskus, Tallinn, Eesti; Teadur
2012 – 2013	Genomatica, Inc, San Diego, Ameerika Ühendriigid; Visiting Scholar "Professional Internship" raames
2007 – 2008	Tokyo University of Electro-Communications, Tokyo, Jaapan; M.Sc. uurimustöö

Täiendõpe

2013	"Advanced Course on Metabolic Engineering and Systems Biology"; Chalmers University of Technology, Rootsi
2012	"Advanced Course on Microbial Physiology and Fermentation Technology"; Delft University of Technology, Holland
2011	FEBS-SystemsX Advanced Lecture Course on Systems Biology; Innsbruck, Austria
2009	"Flow Cytometry course"; University of York, Suurbritannia

Õpetamine ja juhendamine

2014	Karl Peebo M.Sc. töö "Specific growth rate dependent proteome dynamics of *Escherichia coli* in glucose minimal and defined rich medium" juhendaja
2006 – hetkeni	Bakalaureuse tudengite ja magistrantide juhendaja Prof. Raivo Vilu uurimisgrupis Toidu- ja Fermentatsioonitehnoloogia Arenduskeskuses
2010	Lektor ja praktilise kursuse juhendaja "Advanced Microbial Fermentation Technologies", Tallinna Tehnikaülikool

Tunnustused

2012	Baltic-American Freedom Foundation Professional Internship Program stipendium
2011	Artur Linnu stipendium, SA Geenikeskus
2010	II preemia üliõpilaste riikliku teadustööde konkurssi bio- ja keskkonnateaduste valdkonnas, magistriõppe aste
2010	III preemia TTÜ tudengite teadustööde konkurssi loodusteaduste valdkonnas, magistriõppe aste
2009	Jaan Poska stipendium
2007	Kristjan Jaak stipendium

Keelteoskus

Kõrgtase	eesti, inglise
Kesktase	jaapani, soome
Algtase	vene

DISSERTATIONS DEFENDED AT
TALLINN UNIVERSITY OF TECHNOLOGY ON
NATURAL AND EXACT SCIENCES

1. **Olav Kongas**. Nonlinear Dynamics in Modeling Cardiac Arrhytmias. 1998.
2. **Kalju Vanatalu**. Optimization of Processes of Microbial Biosynthesis of Isotopically Labeled Biomolecules and Their Complexes. 1999.
3. **Ahto Buldas**. An Algebraic Approach to the Structure of Graphs. 1999.
4. **Monika Drews**. A Metabolic Study of Insect Cells in Batch and Continuous Culture: Application of Chemostat and Turbidostat to the Production of Recombinant Proteins. 1999.
5. **Eola Valdre**. Endothelial-Specific Regulation of Vessel Formation: Role of Receptor Tyrosine Kinases. 2000.
6. **Kalju Lott**. Doping and Defect Thermodynamic Equilibrium in ZnS. 2000.
7. **Reet Koljak**. Novel Fatty Acid Dioxygenases from the Corals *Plexaura homomalla* and *Gersemia fruticosa.* 2001.
8. **Anne Paju**. Asymmetric oxidation of Prochiral and Racemic Ketones by Using Sharpless Catalyst. 2001.
9. **Marko Vendelin**. Cardiac Mechanoenergetics *in silico*. 2001.
10. **Pearu Peterson**. Multi-Soliton Interactions and the Inverse Problem of Wave Crest. 2001.
11. **Anne Menert**. Microcalorimetry of Anaerobic Digestion. 2001.
12. **Toomas Tiivel**. The Role of the Mitochondrial Outer Membrane in *in vivo* Regulation of Respiration in Normal Heart and Skeletal Muscle Cell. 2002.
13. **Olle Hints**. Ordovician Scolecodonts of Estonia and Neighbouring Areas: Taxonomy, Distribution, Palaeoecology, and Application. 2002.
14. **Jaak Nõlvak**. Chitinozoan Biostratigrapy in the Ordovician of Baltoscandia. 2002.
15. **Liivi Kluge**. On Algebraic Structure of Pre-Operad. 2002.
16. **Jaanus Lass**. Biosignal Interpretation: Study of Cardiac Arrhytmias and Electromagnetic Field Effects on Human Nervous System. 2002.
17. **Janek Peterson**. Synthesis, Structural Characterization and Modification of PAMAM Dendrimers. 2002.
18. **Merike Vaher**. Room Temperature Ionic Liquids as Background Electrolyte Additives in Capillary Electrophoresis. 2002.
19. **Valdek Mikli**. Electron Microscopy and Image Analysis Study of Powdered Hardmetal Materials and Optoelectronic Thin Films. 2003.
20. **Mart Viljus**. The Microstructure and Properties of Fine-Grained Cermets. 2003.
21. **Signe Kask**. Identification and Characterization of Dairy-Related *Lactobacillus.* 2003
22. **Tiiu-Mai Laht**. Influence of Microstructure of the Curd on Enzymatic and Microbiological Processes in Swiss-Type Cheese. 2003.
23. **Anne Kuusksalu**. 2–5A Synthetase in the Marine Sponge *Geodia cydonium*. 2003.
24. **Sergei Bereznev**. Solar Cells Based on Polycristalline Copper-Indium Chalcogenides and Conductive Polymers. 2003.

25. **Kadri Kriis**. Asymmetric Synthesis of C_2-Symmetric Bimorpholines and Their Application as Chiral Ligands in the Transfer Hydrogenation of Aromatic Ketones. 2004.
26. **Jekaterina Reut**. Polypyrrole Coatings on Conducting and Insulating Substracts. 2004.
27. **Sven Nõmm**. Realization and Identification of Discrete-Time Nonlinear Systems. 2004.
28. **Olga Kijatkina**. Deposition of Copper Indium Disulphide Films by Chemical Spray Pyrolysis. 2004.
29. **Gert Tamberg**. On Sampling Operators Defined by Rogosinski, Hann and Blackman Windows. 2004.
30. **Monika Übner**. Interaction of Humic Substances with Metal Cations. 2004.
31. **Kaarel Adamberg**. Growth Characteristics of Non-Starter Lactic Acid Bacteria from Cheese. 2004.
32. **Imre Vallikivi**. Lipase-Catalysed Reactions of Prostaglandins. 2004.
33. **Merike Peld**. Substituted Apatites as Sorbents for Heavy Metals. 2005.
34. **Vitali Syritski**. Study of Synthesis and Redox Switching of Polypyrrole and Poly(3,4-ethylenedioxythiophene) by Using *in-situ* Techniques. 2004.
35. **Lee Põllumaa**. Evaluation of Ecotoxicological Effects Related to Oil Shale Industry. 2004.
36. **Riina Aav**. Synthesis of 9,11-Secosterols Intermediates. 2005.
37. **Andres Braunbrück**. Wave Interaction in Weakly Inhomogeneous Materials. 2005.
38. **Robert Kitt**. Generalised Scale-Invariance in Financial Time Series. 2005.
39. **Juss Pavelson**. Mesoscale Physical Processes and the Related Impact on the Summer Nutrient Fields and Phytoplankton Blooms in the Western Gulf of Finland. 2005.
40. **Olari Ilison**. Solitons and Solitary Waves in Media with Higher Order Dispersive and Nonlinear Effects. 2005.
41. **Maksim Säkki**. Intermittency and Long-Range Structurization of Heart Rate. 2005.
42. **Enli Kiipli**. Modelling Seawater Chemistry of the East Baltic Basin in the Late Ordovician–Early Silurian. 2005.
43. **Igor Golovtsov**. Modification of Conductive Properties and Processability of Polyparaphenylene, Polypyrrole and polyaniline. 2005.
44. **Katrin Laos**. Interaction Between Furcellaran and the Globular Proteins (Bovine Serum Albumin β-Lactoglobulin). 2005.
45. **Arvo Mere**. Structural and Electrical Properties of Spray Deposited Copper Indium Disulphide Films for Solar Cells. 2006.
46. **Sille Ehala**. Development and Application of Various On- and Off-Line Analytical Methods for the Analysis of Bioactive Compounds. 2006.
47. **Maria Kulp**. Capillary Electrophoretic Monitoring of Biochemical Reaction Kinetics. 2006.
48. **Anu Aaspõllu.** Proteinases from *Vipera lebetina* Snake Venom Affecting Hemostasis. 2006.

49. **Lyudmila Chekulayeva**. Photosensitized Inactivation of Tumor Cells by Porphyrins and Chlorins. 2006.
50. **Merle Uudsemaa**. Quantum-Chemical Modeling of Solvated First Row Transition Metal Ions. 2006.
51. **Tagli Pitsi**. Nutrition Situation of Pre-School Children in Estonia from 1995 to 2004. 2006.
52. **Angela Ivask**. Luminescent Recombinant Sensor Bacteria for the Analysis of Bioavailable Heavy Metals. 2006.
53. **Tiina Lõugas**. Study on Physico-Chemical Properties and Some Bioactive Compounds of Sea Buckthorn (*Hippophae rhamnoides* L.). 2006.
54. **Kaja Kasemets**. Effect of Changing Environmental Conditions on the Fermentative Growth of *Saccharomyces cerevisae* S288C: Auxo-accelerostat Study. 2006.
55. **Ildar Nisamedtinov**. Application of ^{13}C and Fluorescence Labeling in Metabolic Studies of *Saccharomyces* spp. 2006.
56. **Alar Leibak**. On Additive Generalisation of Voronoï's Theory of Perfect Forms over Algebraic Number Fields. 2006.
57. **Andri Jagomägi**. Photoluminescence of Chalcopyrite Tellurides. 2006.
58. **Tõnu Martma**. Application of Carbon Isotopes to the Study of the Ordovician and Silurian of the Baltic. 2006.
59. **Marit Kauk**. Chemical Composition of $CuInSe_2$ Monograin Powders for Solar Cell Application. 2006.
60. **Julia Kois**. Electrochemical Deposition of $CuInSe_2$ Thin Films for Photovoltaic Applications. 2006.
61. **Ilona Oja Açik**. Sol-Gel Deposition of Titanium Dioxide Films. 2007.
62. **Tiia Anmann**. Integrated and Organized Cellular Bioenergetic Systems in Heart and Brain. 2007.
63. **Katrin Trummal**. Purification, Characterization and Specificity Studies of Metalloproteinases from *Vipera lebetina* Snake Venom. 2007.
64. **Gennadi Lessin**. Biochemical Definition of Coastal Zone Using Numerical Modeling and Measurement Data. 2007.
65. **Enno Pais**. Inverse problems to determine non-homogeneous degenerate memory kernels in heat flow. 2007.
66. **Maria Borissova**. Capillary Electrophoresis on Alkylimidazolium Salts. 2007.
67. **Karin Valmsen**. Prostaglandin Synthesis in the Coral *Plexaura homomalla*: Control of Prostaglandin Stereochemistry at Carbon 15 by Cyclooxygenases. 2007.
68. **Kristjan Piirimäe**. Long-Term Changes of Nutrient Fluxes in the Drainage Basin of the Gulf of Finland – Application of the PolFlow Model. 2007.
69. **Tatjana Dedova**. Chemical Spray Pyrolysis Deposition of Zinc Sulfide Thin Films and Zinc Oxide Nanostructured Layers. 2007.
70. **Katrin Tomson**. Production of Labelled Recombinant Proteins in Fed-Batch Systems in *Escherichia coli*. 2007.
71. **Cecilia Sarmiento**. Suppressors of RNA Silencing in Plants. 2008.

72. **Vilja Mardla**. Inhibition of Platelet Aggregation with Combination of Antiplatelet Agents. 2008.
73. **Maie Bachmann**. Effect of Modulated Microwave Radiation on Human Resting Electroencephalographic Signal. 2008.
74. **Dan Hüvonen**. Terahertz Spectroscopy of Low-Dimensional Spin Systems. 2008.
75. **Ly Villo**. Stereoselective Chemoenzymatic Synthesis of Deoxy Sugar Esters Involving *Candida antarctica* Lipase B. 2008.
76. **Johan Anton**. Technology of Integrated Photoelasticity for Residual Stress Measurement in Glass Articles of Axisymmetric Shape. 2008.
77. **Olga Volobujeva**. SEM Study of Selenization of Different Thin Metallic Films. 2008.
78. **Artur Jõgi**. Synthesis of 4'-Substituted 2,3'-dideoxynucleoside Analogues. 2008.
79. **Mario Kadastik**. Doubly Charged Higgs Boson Decays and Implications on Neutrino Physics. 2008.
80. **Fernando Pérez-Caballero**. Carbon Aerogels from 5-Methylresorcinol-Formaldehyde Gels. 2008.
81. **Sirje Vaask**. The Comparability, Reproducibility and Validity of Estonian Food Consumption Surveys. 2008.
82. **Anna Menaker**. Electrosynthesized Conducting Polymers, Polypyrrole and Poly(3,4-ethylenedioxythiophene), for Molecular Imprinting. 2009.
83. **Lauri Ilison**. Solitons and Solitary Waves in Hierarchical Korteweg-de Vries Type Systems. 2009.
84. **Kaia Ernits**. Study of In_2S_3 and ZnS Thin Films Deposited by Ultrasonic Spray Pyrolysis and Chemical Deposition. 2009.
85. **Veljo Sinivee**. Portable Spectrometer for Ionizing Radiation "Gammamapper". 2009.
86. **Jüri Virkepu**. On Lagrange Formalism for Lie Theory and Operadic Harmonic Oscillator in Low Dimensions. 2009.
87. **Marko Piirsoo**. Deciphering Molecular Basis of Schwann Cell Development. 2009.
88. **Kati Helmja**. Determination of Phenolic Compounds and Their Antioxidative Capability in Plant Extracts. 2010.
89. **Merike Sõmera**. Sobemoviruses: Genomic Organization, Potential for Recombination and Necessity of P1 in Systemic Infection. 2010.
90. **Kristjan Laes**. Preparation and Impedance Spectroscopy of Hybrid Structures Based on $CuIn_3Se_5$ Photoabsorber. 2010.
91. **Kristin Lippur**. Asymmetric Synthesis of 2,2'-Bimorpholine and its 5,5'-Substituted Derivatives. 2010.
92. **Merike Luman**. Dialysis Dose and Nutrition Assessment by an Optical Method. 2010.
93. **Mihhail Berezovski**. Numerical Simulation of Wave Propagation in Heterogeneous and Microstructured Materials. 2010.
94. **Tamara Aid-Pavlidis**. Structure and Regulation of BDNF Gene. 2010.
95. **Olga Bragina**. The Role of Sonic Hedgehog Pathway in Neuro- and Tumorigenesis. 2010.

96. **Merle Randrüüt**. Wave Propagation in Microstructured Solids: Solitary and Periodic Waves. 2010.
97. **Marju Laars**. Asymmetric Organocatalytic Michael and Aldol Reactions Mediated by Cyclic Amines. 2010.
98. **Maarja Grossberg**. Optical Properties of Multinary Semiconductor Compounds for Photovoltaic Applications. 2010.
99. **Alla Maloverjan**. Vertebrate Homologues of Drosophila Fused Kinase and Their Role in Sonic Hedgehog Signalling Pathway. 2010.
100. **Priit Pruunsild**. Neuronal Activity-Dependent Transcription Factors and Regulation of Human *BDNF* Gene. 2010.
101. **Tatjana Knjazeva**. New Approaches in Capillary Electrophoresis for Separation and Study of Proteins. 2011.
102. **Atanas Katerski**. Chemical Composition of Sprayed Copper Indium Disulfide Films for Nanostructured Solar Cells. 2011.
103. **Kristi Timmo.** Formation of Properties of $CuInSe_2$ and $Cu_2ZnSn(S,Se)_4$ Monograin Powders Synthesized in Molten KI. 2011.
104. **Kert Tamm**. Wave Propagation and Interaction in Mindlin-Type Microstructured Solids: Numerical Simulation. 2011.
105. **Adrian Popp**. Ordovician Proetid Trilobites in Baltoscandia and Germany. 2011.
106. **Ove Pärn**. Sea Ice Deformation Events in the Gulf of Finland and This Impact on Shipping. 2011.
107. **Germo Väli**. Numerical Experiments on Matter Transport in the Baltic Sea. 2011.
108. **Andrus Seiman**. Point-of-Care Analyser Based on Capillary Electrophoresis. 2011.
109. **Olga Katargina**. Tick-Borne Pathogens Circulating in Estonia (Tick-Borne Encephalitis Virus, *Anaplasma phagocytophilum*, *Babesia* Species): Their Prevalence and Genetic Characterization. 2011.
110. **Ingrid Sumeri**. The Study of Probiotic Bacteria in Human Gastrointestinal Tract Simulator. 2011.
111. **Kairit Zovo**. Functional Characterization of Cellular Copper Proteome. 2011.
112. **Natalja Makarytsheva**. Analysis of Organic Species in Sediments and Soil by High Performance Separation Methods. 2011.
113. **Monika Mortimer**. Evaluation of the Biological Effects of Engineered Nanoparticles on Unicellular Pro- and Eukaryotic Organisms. 2011.
114. **Kersti Tepp**. Molecular System Bioenergetics of Cardiac Cells: Quantitative Analysis of Structure-Function Relationship. 2011.
115. **Anna-Liisa Peikolainen**. Organic Aerogels Based on 5-Methylresorcinol. 2011.
116. **Leeli Amon**. Palaeoecological Reconstruction of Late-Glacial Vegetation Dynamics in Eastern Baltic Area: A View Based on Plant Macrofossil Analysis. 2011.
117. **Tanel Peets**. Dispersion Analysis of Wave Motion in Microstructured Solids. 2011.
118. **Liina Kaupmees**. Selenization of Molybdenum as Contact Material in Solar Cells. 2011.
119. **Allan Olspert**. Properties of VPg and Coat Protein of Sobemoviruses. 2011.

120. **Kadri Koppel**. Food Category Appraisal Using Sensory Methods. 2011.
121. **Jelena Gorbatšova**. Development of Methods for CE Analysis of Plant Phenolics and Vitamins. 2011.
122. **Karin Viipsi**. Impact of EDTA and Humic Substances on the Removal of Cd and Zn from Aqueous Solutions by Apatite. 2012.
123. **David Schryer**. Metabolic Flux Analysis of Compartmentalized Systems Using Dynamic Isotopologue Modeling. 2012.
124. **Ardo Illaste**. Analysis of Molecular Movements in Cardiac Myocytes. 2012.
125. **Indrek Reile**. 3-Alkylcyclopentane-1,2-Diones in Asymmetric Oxidation and Alkylation Reactions. 2012.
126. **Tatjana Tamberg**. Some Classes of Finite 2-Groups and Their Endomorphism Semigroups. 2012.
127. **Taavi Liblik**. Variability of Thermohaline Structure in the Gulf of Finland in Summer. 2012.
128. **Priidik Lagemaa**. Operational Forecasting in Estonian Marine Waters. 2012.
129. **Andrei Errapart**. Photoelastic Tomography in Linear and Non-linear Approximation. 2012.
130. **Külliki Krabbi**. Biochemical Diagnosis of Classical Galactosemia and Mucopolysaccharidoses in Estonia. 2012.
131. **Kristel Kaseleht**. Identification of Aroma Compounds in Food using SPME-GC/MS and GC-Olfactometry. 2012.
132. **Kristel Kodar**. Immunoglobulin G Glycosylation Profiling in Patients with Gastric Cancer. 2012.
133. **Kai Rosin**. Solar Radiation and Wind as Agents of the Formation of the Radiation Regime in Water Bodies. 2012.
134. **Ann Tiiman**. Interactions of Alzheimer's Amyloid-Beta Peptides with Zn(II) and Cu(II) Ions. 2012.
135. **Olga Gavrilova**. Application and Elaboration of Accounting Approaches for Sustainable Development. 2012.
136. **Olesja Bondarenko**. Development of Bacterial Biosensors and Human Stem Cell-Based *In Vitro* Assays for the Toxicological Profiling of Synthetic Nanoparticles. 2012.
137. **Katri Muska**. Study of Composition and Thermal Treatments of Quaternary Compounds for Monograin Layer Solar Cells. 2012.
138. **Ranno Nahku**. Validation of Critical Factors for the Quantitative Characterization of Bacterial Physiology in Accelerostat Cultures. 2012.
139. **Petri-Jaan Lahtvee**. Quantitative Omics-level Analysis of Growth Rate Dependent Energy Metabolism in *Lactococcus lactis*. 2012.
140. **Kerti Orumets**. Molecular Mechanisms Controlling Intracellular Glutathione Levels in Baker's Yeast *Saccharomyces cerevisiae* and its Random Mutagenized Gluthatione Over-Accumulating Isolate. 2012.
141. **Loreida Timberg**. Spice-Cured Sprats Ripening, Sensory Parameters Development, and Quality Indicators. 2012.
142. **Anna Mihhalevski**. Rye Sourdough Fermentation and Bread Stability. 2012.

143. **Liisa Arike**. Quantitative Proteomics of *Escherichia coli*: From Relative to Absolute Scale. 2012.
144. **Kairi Otto**. Deposition of In_2S_3 Thin Films by Chemical Spray Pyrolysis. 2012.
145. **Mari Sepp**. Functions of the Basic Helix-Loop-Helix Transcription Factor TCF4 in Health and Disease. 2012.
146. **Anna Suhhova**. Detection of the Effect of Weak Stressors on Human Resting Electroencephalographic Signal. 2012.
147. **Aram Kazarjan**. Development and Production of Extruded Food and Feed Products Containing Probiotic Microorganisms. 2012.
148. **Rivo Uiboupin**. Application of Remote Sensing Methods for the Investigation of Spatio-Temporal Variability of Sea Surface Temperature and Chlorophyll Fields in the Gulf of Finland. 2013.
149. **Tiina Kriščiunaite**. A Study of Milk Coagulability. 2013.
150. **Tuuli Levandi**. Comparative Study of Cereal Varieties by Analytical Separation Methods and Chemometrics. 2013.
151. **Natalja Kabanova**. Development of a Microcalorimetric Method for the Study of Fermentation Processes. 2013.
152. **Himani Khanduri**. Magnetic Properties of Functional Oxides. 2013.
153. **Julia Smirnova**. Investigation of Properties and Reaction Mechanisms of Redox-Active Proteins by ESI MS. 2013.
154. **Mervi Sepp**. Estimation of Diffusion Restrictions in Cardiomyocytes Using Kinetic Measurements. 2013.
155. **Kersti Jääger**. Differentiation and Heterogeneity of Mesenchymal Stem Cells. 2013.
156. **Victor Alari**. Multi-Scale Wind Wave Modeling in the Baltic Sea. 2013.
157. **Taavi Päll**. Studies of CD44 Hyaluronan Binding Domain as Novel Angiogenesis Inhibitor. 2013.
158. **Allan Niidu**. Synthesis of Cyclopentane and Tetrahydrofuran Derivatives. 2013.
159. **Julia Geller**. Detection and Genetic Characterization of *Borrelia* Species Circulating in Tick Population in Estonia. 2013.
160. **Irina Stulova**. The Effects of Milk Composition and Treatment on the Growth of Lactic Acid Bacteria. 2013.
161. **Jana Holmar**. Optical Method for Uric Acid Removal Assessment During Dialysis. 2013.
162. **Kerti Ausmees**. Synthesis of Heterobicyclo[3.2.0]heptane Derivatives *via* Multicomponent Cascade Reaction. 2013.
163. **Minna Varikmaa**. Structural and Functional Studies of Mitochondrial Respiration Regulation in Muscle Cells. 2013.
164. **Indrek Koppel**. Transcriptional Mechanisms of BDNF Gene Regulation. 2014.
165. **Kristjan Pilt**. Optical Pulse Wave Signal Analysis for Determination of Early Arterial Ageing in Diabetic Patients. 2014.
166. **Andres Anier**. Estimation of the Complexity of the Electroencephalogram for Brain Monitoring in Intensive Care. 2014.

167. **Toivo Kallaste**. Pyroclastic Sanidine in the Lower Palaeozoic Bentonites – A Tool for Regional Geological Correlations. 2014.
168. **Erki Kärber**. Properties of ZnO-nanorod/In_2S_3/$CuInS_2$ Solar Cell and the Constituent Layers Deposited by Chemical Spray Method. 2014.
169. **Julia Lehner**. Formation of Cu_2ZnSnS_4 and $Cu_2ZnSnSe_4$ by Chalcogenisation of Electrochemically Deposited Precursor Layers. 2014.