LC-MS based Metabolomics

Magdenoska, Olivera; Nielsen, Kristian Fog; Thykær, Jette

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

LC-MS BASED METABOLOMICS

Olivera Magdenoska

PhD Thesis

Technical University of Denmark

Department of Systems Biology

Eukaryotic Biotechnology

Preface

This PhD study was carried out at the section for Eukaryotic Biotechnology, Department of Systems Biology, Technical University of Denmark in the period of 1st of July 2011 to 30th of April 2015. This work was partly supported by the Danish Research Agency for Technology and Production, grant # 09-064967.

First and foremost I would like to thank my supervisor ***Kristian Fog Nielsen*** for his excellent supervision. Without his expertise, thoughtful guidance and scientific discussions this work would not have been possible. I would also like to thank my co-supervisor ***Jette Thykær*** for her assistance, advices and constructive feedback when most needed. Thank you for offering your comprehensive knowledge.

Many thanks to my office mate and great friend ***Ina Došen.*** Thank you for being there whenever I needed a friend. I will really miss our entertaining discussions and long laughs. A special thanks goes to ***Anna Lena Heins*****.** Having your support and friendship has meant a lot during good times and those when things haven't gone so well. I would also like to thank my previous office mates, ***Anita Iversen*** and ***Yuksel Gezgin***. It has been a pleasure to work with you.

I would also like to thank my colleagues ***Peter Boldsen Knudsen***, ***Subir Kumar Nandy, Paul W. D'Alvise, Tomas Strucko, Daniel Killerup Svenssen*** for the good team work.

A big thanks goes also to all the people from ''CMB'' for creating relaxing working atmosphere. I have had 4 amazing years filled with laughs, social events and will therefore miss you all.

This journey would not have been accomplished without ***Gustav Hammerich Hansen***. Thank you for being endlessly patient, kind, loving and supportive. Finally I take this opportunity to express my deep gratitude to my beloved ***Parents*** and my ***Sister***, for their precious support and unconditional love.

Summary

Metabolomics, the qualitative and quantitative analysis of metabolites is a valuable approach for understanding the biochemical processes in the cells. Particularly important are the intracellular metabolites that supply the cell with energy, serve as building blocks and act as signaling molecules. The analytical tools applied for analysis of intracellular metabolites should be capable to cope with the large number of metabolites to be analyzed and the complex matrix in the samples. Therefore the combination of separation and detection techniques is commonly applied for analysis of intracellular metabolites with liquid chromatography mass spectrometry (LC-MS) as the most commonly used.

The primary goal of this Ph.D. study was to develop an LC-MS method together with sample preparation for analysis of intracellular metabolites such as nucleotides, sugar phosphates, organic acids, coenzymes etc.. In the studies conducted during this Ph.D. the developed method was used to understand how the genetic manipulations in various organisms, influence the levels of their intracellular metabolites. The method development was divided into three steps: i) optimization of the MS detection, ii) establishment and optimization of the chromatographic separation and iii) optimization of the sample preparation. A substantial part of the thesis was focused on the development of the LC-MS method. For quantitative targeted analysis of a group of defined metabolites, triple quadrupole (QqQ) MS was used. The optimization of the MS detection aimed to determine multiple reaction monitoring (MRM) transitions of the analytes and to increase the sensitivity by testing different ion-source parameters and collision energies. This resulted in optimized detection of more than 50 intracellular metabolites. During the optimization of the chromatographic separation, anion exchange (AEC), ion chromatography (IC) and ion-pair reversed phase (IP-RP) were tested with the ion-pair giving the best compromise between retention, separation and stability of the compounds during the chromatographic separation. By testing different types and

concentrations of ion-pair reagent and different concentrations of acetic acid as a counter ion, it was found that a solution of 10 mM tributylamine (TBA) and 10 mM acetic acid gave the best compromise between chromatographic retention and separation.

Establishment of proper sample preparation (quenching and extraction) procedures for intracellular metabolites was necessary in order to obtain meaningful metabolomics data. The main idea was to find a sample preparation method that will give the best compromise between an acceptable energy charge ratio (ECR, usually between 0.80-0.95) low leakage during the quenching and high recovery of the metabolites after the extraction. Quenching and extraction procedures for bacteria, yeast, mammalian cells and filamentous fungi were tested. Cold MeOH as a quenching method combined with boiling EtOH or MeOH/chloroform as extraction method showed to work well for *Saccharomyces cerevisiae* (*S. cerevisiae*) resulting in an ECR of 0.80-0.95 and less than 10 % leakage. Quenching of bacteria and fungi showed to be challenging task due to the high susceptibility of these organisms to leakage during quenching. Quenching using formic acid, where the cells were not separated from the media, was shown to work well for *Lactococcus lactis* (*L. lactis*) but not for *Streptomyces coelicolor* (*S. coelicolor*) and *Microbispora corallina* (*M. corallina*). The reason for this was speculated to be due to the filamentous growth of these organisms. Other quenching procedures were tested for *S. coelicolor* and *M. corallina* with -40 °C MeOH/H_2O (60/40, v/v) giving an acceptable ECR (in the range of 0.80-0.95). However leakage was observed for both organisms. For filamentous fungi, filtration in combination with cold methanol or 0 °C saline resulted in successful quenching. In the case of cold methanol quenching the concentration of AMP and ADP in the quenching supernatant was found to be 30 % of the total amount found in the biomass and the supernatant. Saline at 0 °C showed to be a good quenching solution for mammalian cells as well and was combined with an extraction procedure based on the addition of cold MeOH and ACN/H_2O (50/50, v/v).

In another study conducted during this Ph.D. a novel approach for creating an authentic matrix that can be used for validation of analytical methods was established. This circumvents the problem with the absence of a matrix free of the analyte which is needed for preparation of the calibration curves for quantification of intracellular metabolites. The spiking matrix was produced by extracting biomass obtained from growing *S. cerevisiae* in a media that contained ^{13}C labeled glucose/non-labeled glucose (50/50, w/w). The advantage of this matrix was that the pools of the compounds with only ^{12}C or ^{13}C carbons were very low or even not measurable and showed minimal or no interference to the spiked amount of non-labeled standards and their stable isotope-labeled internal standards (SIL-IS).

Finally the developed IP-RP LC-MS method was coupled to a quadrupole time of flight (Q-TOF) MS for multitargeted analysis, detection and identification of as many known unknown metabolites as possible in different biological matrices. The feasibility of using the Q-TOF MS was evaluated by analysis of extracts from three different organisms: *S. cerevisiae*, *M. corallina* and *S. coelicolor*. The screening concept in this study was based on two approaches: i) aggressive dereplication of the full scan high resolution MS (HR-MS) data using search lists of known compounds and ii) high resolution tandem MS (MS/HRMS) data searched in Metabolite Link (METLIN) library.

The data presented here show that the methods developed during this Ph.D. study were successfully applied for targeted and multitargeted analysis of different classes of intracellular metabolites such as nucleotides, sugar phosphates, coeznymes and organic acids. In addition sample preparation methods were established for different microorganisms capable of extracting broad range of metabolites. Finally these methods have shown to be valuable addition to the ‘’omics’’ tools used to reveal key information regarding the metabolism and the regulation in the biological systems.

Sammenfatning

Metabolomics, den kvalitative og kvantitative analyse af metabolitter er en værdifuld metode til forståelsen af de biokemiske processer i cellerne. Særligt vigtige er de intracellulære metabolitter, der supplerer cellen med energi, bruges som byggeklodser og fungerer som signalmolekyler. De analytiske redskaber anvendt til måling af intracellulære metabolitter bør være i stand til at håndtere den store mængde af metabolitter, der skal analyseres og den komplekse matrix i prøverne. Derfor anvender man normalt kombinationen af separation og detektion, til analysen af intracellulære metabolitter med flydende kromatografi og masse spektrometri (LC-MS) som den mest almindeligt anvendte.

Hovedformålet med dette Ph.D. studium var at udvikle en LC-MS metode sammen med prøve forberedelse til analyse af intracellulære metabolitter såsom nukleotider, sukker fosfater, organiske syre, co-enzymer etc.. I studierne udført under denne Ph.D., blev den udviklede metode brugt til at forstå hvordan den genetiske manipulering i forskellige organismer påvirker niveauet af deres intracellulære metabolitter. Metode udviklingen blev inddelt i tre trin: i) optimering af MS-detektionen, ii) etablering og optimering af den kromatografiske separation og iii) optimering af prøve forberedelse. En betydelig del af afhandlingen fokuserede på udviklingen af denne LC-MS metode. Triple quadrupole (QqQ) MS blev anvendt til kvantitativ målrettet analyse af en gruppe af definerede metabolitter. Optimeringen af MS detektionen stilede imod at bestemme multiple reaction monitoring (MRM) i.f.t. overgange af analytterne og mod at forøge følsomheden ved at teste forskellige ion-kilde parametre samt kollisions energier. Dette resulterede i optimeret detektion af af mere end 50 intracellulære metabolitter. Undervejs i optimeringen af kromatografisk adskillelse blev anion bytning (AEC), ion kromatografi (IC) og ion-par omvendt fase (IP-RP) testet med det ion-par der gav det bedste kompromis mellem retention, adskillelse og stabilitet af stofferne igennem den kromatografiske adskillelse. Ved testning af forskellige

typer og koncentrationer af ion-par reagent og forskellige koncentrationer af eddikesyre som en modbalance ion blev det vist at en opløsning af 10 mM tributylamine og 10 mM eddikesyre gav det bedste kompromis mellem kromatografisk retention og adskillelse.

Etableringen af ordentlige prøve forberedelse (quenching og ekstraktion) procedure til intracellulære metabolitter var nødvendig for at kunne opnå betydningsfulde data. Hovedformålet var at finde en metode, der ville give det bedste kompromis mellem et acceptabelt energi ladnings forhold (ECR, normal mellem 0.80-0.95), lav lækage under quenchingen og høj gendannelse af metabolitterne efter ekstraktionen. Quenching og ekstraktions procedure for bakterier, gær, mammale celler og filamentiøse svampe blev analyseret. Quenchingen af bakterier og svampe viste sig at være en udfordrende opgave på grund af den høje modtagelighed af disse organismer overfor lækage under quenchingen. Quenching med myresyre, hvor cellerne ikke blev adskilt fra mediet, blev påvist at virke godt for *Lactococcus lactis* (*L. lactis*), men ikke for *Streptomyces coelicolor* (*S. coelicolor*) eller *Microbispora corallina* (*M. corallina*). Årsagen til dette blev antaget at være på grund af organismernes filamentiøse vækst. Andre quenching procedurer blev testet for *S. coelicolor* og *M. corallina* med -40 °C MeOH/H_2O (60/40, v/v), resulterende i en acceptabel ECR (i området 0.80-0.90). Dog blev lækage observeret for begge organismer. For filamentiøse svampe blev filtrering i kombination med kold metanol brugt til quenching. Men koncentrationerne af AMP og ADP i quenching supernatanten var omkring 30 % af den totale mængde fundet i biomassen og supernatanten. Kold MeOH som en quenching metode i kombination med EtOH eller MeOH/kloroform som ekstraktions metode viste sig at fungere godt for *Saccharomyces cerevisiae* (*S. cerevisiae*). Saltvand ved 0 °C viste sig, som en god quenchings opløsning for mammale celler imens ekstraktions proceduren var baseret på tilsætning af kold MeOH og ACN/H_2O (50/50, v/v).

I et andet studie udført under denne Ph.D. blev en ny tilgang etableret til dannelsen af en kunstig matrix som kan bruges til validering af analytiske metoder. Med dette omgås problemet med mangel af en matrix fri for analytterne, som er nødvendig til forberedelsen af kalibreringskurverne til kvantificering af intracellulære metabolitter. Spiking matricen blev produceret ved at ekstrahere biomasse opnået fra voksende *S. cerevisiae* i et medie, som indeholdt ^{13}C mærket glukose/umærket glukose (50/50, w/w). Fordelen ved at bruge denne matrix var, at puljerne af stoffer med kun ^{12}C eller ^{13}C kulstoffer var meget lave eller næsten ikke målbare og viste minimal eller ingen interferens af den spikede mængde af umærkede standarder og deres stabile isotopiske mærkede interne standarder (SIL-IS).

Til slut blev den udviklede IR-RP LC-MS metode koblet til en quadropol time of flight (Q-TOF) MS til multimålrettet analyse, detektion og identifikation af så mange kendte ukendte metabolitter, som muligt i forskellige biologiske matricer. Gennemførligheden ved brugen af Q-TOF MS blev evalueret ved at analysere ekstrakter fra tre forskellige organismer: *S. cerevisiae*, *M. corallina* og *S. coelicolor*. Screenings konceptet i dette studie, var baseret på to tilgange: i) aggresiv dereplikation af fuldt scannet høj opløsning MS (HR-MS) data ved brug af søgningslister med kendte stoffer og ii) høj opløsning tandem MS (MS/HRMS) data gennemsøgt i Metabolite Link (METLIN) biblioteket.

Den data, som er blevet præsenteret her viser at metoderne udviklet under dette Ph.D. studium blev succesfuldt anvendt til målrettet og multimålrettet analyse af forskellige klasser af intracellulære metabolitter såsom nukleotider, sukker fosfater, coenzymer og organiske syre. Endvidere blev prøve forberedelses metoder etableret til forskellige mikroorganismer der let kan ekstrahere et bredt udvalg af metabolitter med god gendannelse. Slutteligt blev disse metoder påvist at være en værdifuld tilføjelse til ”omics” værktøjerne brugt til at afslører nøgle informationer angående metabolismen og reguleringen i biologiske systemer.

List of publications and other communications

Paper 1 **Olivera Magdenoska**; Jan Martinussen; Jette Thykær; Kristian Fog Nielsen, Dispersive solid phase extraction combined with ion-pair ultra high-performance liquid chromatography tandem mass spectrometry for quantification of nucleotides in *Lactococcus lactis*. Anal. Biochem., 440 **(2013)**166-177. (**Published**)

Paper 2 Paul W. D'Alvise; **Olivera Magdenoska**; Jette Melchiorsen; Kristian Fog Nielsen; Lone Gram, Biofilm formation and antibiotic production in *Ruegeria mobilis* are influenced by intracellular concentrations of cyclic dimeric guanosin monophosphate, Environ. Microbiol., 16 **(2014)** 1252-1266. (**Published**)

Paper 3 Daniel Ley, Ali Kazemi Seresht, Mikael Engmark, **Olivera Magdenoska**, Kristian Fog Nielsen, Helene Faustrup Kildegaard, Mikael Rørdam Andersen, Multi-omic profiling of EPO-producing Chinese hamster ovary cell panel reveals metabolic adaptation to heterologous protein production, Biotechnol. Bioeng. (**Submitted**)

Paper 4 Tomas Strucko, **Olivera Magdenoska**, Uffe Mortensen, Benchmarking two commonly used Saccharomyces cerevisiae strains for heterologous vanillin-β-glucoside production, Metab. Eng. Commun. (**Submitted**)

Paper 5 **Olivera Magdenoska,** Peter Boldsen Knudsen, Daniel Killerup Svenssen, Kristian Fog Nielsen. LC-MS/MS quantification of intracellular metabolites in Saccharomyces cerevisiae using ^{13}C-labeling to minimize matrix interference. Anal. Biochem. **(Submitted**)

Paper 6 Multitargeted analysis of intracellular metabolites in various microorganisms using ion-pair reversed phase UHPLC-Q-TOF MS (**Manuscript in preparation**)

Paper 7 **Olivera Magdenoska**, Daniel Killerup Svenssen, Peter Boldsen Knudsen, Andrea Thorhallsdottir, Mhairi Workman og Kristian Fog Nielsen Metabolomets ioniske komponenter bestemt ved kromatografi og massespektrometri, Dansk Kemi, 96 nr.5, 2015. (**Published non-peer reviewed**)

Poster communications:

Evaluation of cell factory performance through determination of intracellular metabolites using LC-MS/MS

Magdenoska, Olivera; Martinussen, Jan; Nielsen, Kristian Fog; Thykær, Jette in journal: New Biotechnology (ISSN: 1871-6784) (DOI: http://dx.doi.org/10.1016/j.nbt.2012.08.413), vol: 29S, 2012

Presented at: 15th European Congress on Biotechnology, Istanbul

Type: Conference abstract in journal (Peer reviewed)

Status: Published | Year: 2012 | DOI: http://dx.doi.org/10.1016/j.nbt.2012.08.413

Ion-pair UHPLC-QTOFMS for intracellular metabolomics of various microorganisms

Magdenoska, Olivera; Nandy, Subir Kumar; Lantz, Anna Eliasson; Svensen, Daniel Killerup; Thykær Jette, Nielsen, Kristian Fog

Presented at Metabolomics 2014 in Tsuruoka, Japan.

Oral communications:

UHPLC-MS/MS target-metabolomics for highly polar and ionic analytes

Magdenoska, Olivera; Thykær, Jette; Nielsen, Kristian Fog

Presented at: 3rd Danish Symposium on Metabolomics 14th of November 2011 in Copenhagen, Denmark

UHPLC-MS/MS target-metabolomics for highly polar and ionic analytes

Magdenoska, Olivera; Thykær, Jette; Nielsen, Kristian Fog

Presented at: The yearly meeting in the Danish Society for Mass Spectrometry, January 19-20, 2012 in Svendborg, Denmark

LC-MS/MS target-metabolomics for highly polar and ionic analytes

Magdenoska, Olivera; Thykær, Jette; Nielsen, Kristian Fog

Presented at: Agilent Nordic Users Meeting, April 19-20, 2012 in Gothenburg, Sweden

Ion-pair UHPLC-MS/MS and Q-TOF for analysis of intracellular metabolites in microorganisms

Magdenoska, Olivera; Thykær, Jette; Nielsen, Kristian Fog

Presented at: Agilent chemical analysis MS meeting and workshops, September 17-19, 2013, Birmingham, England

Abbreviations

2PG	2-Phosphogluceric acid
3PG	3-Phosphogluceric acid
6PG	6-phosphogluconate
Ac-CoA	Acetyl coenzyme A
ACT	Acinorhodin
ADP	Adenosine 5'-diphosphate
AEC	Anion exchange chromatography
AMP	Adenosine 5'-monophosphate
APCI	Atmospheric pressure chemical ionization
ATP	Adenosine 5'-triphosphate
cAMP	Adenosine 3',5'-cyclic phosphate
c-diAMP	Cyclic di-3',5'-adenylate
c-diGMP	Cyclic di-3',5'-guanylate
CDP	Cytidine 5'-diphosphate
CE	Capillary electrophoresis
cGMP	Guanosine 3',5'-cyclic phosphate
CHO	Chinese hamster ovary cells
CMP	Cytidine 5'-monophosphate
CTP	Cytidine 5'-triphosphate
C-VG	Vanillin-β-glucoside producing CEN.PK
dATP	Deoxyadenosine 5'-triphosphate
DBA	Dibutylamine
dCMP	Deoxycytidine 5'-monopdosphate
dCTP	Deoxycytidine 5'-triphosphate
dGMP	Deoxynuanosine 5'-monophosphate
dGTP	Deoxyguanosine 5'-triphosphate
DHAP	Dihydroxyacetone phosphate
dTMP	Deoxythimidine 5'-monophosphate
dUMP	Deoxyuridine 5'-monophosphate
E. coli	*Escherichia coli*
E4P	Erythrose 4-phosphate
ECR	Energy charge ratio
EI	Electron ionization
EMP	Embden-Meyerhof-Parnas
ESI	Electrospray ionization

F1P	Fructose 1-phosphate
F6P	Fructose 6-phosphate
FAD	Flavin adenine dinucleotide
FBP	Fructose 1,6-bisphosphate
G1P	Glucose 1-phosphate
G3P	Glyceraldehyde 3-phosphate
G6P	Glucose 6-phosphate
Gal1P	Galactose 1-phosphate
GC	Gas chromatography
GDP	Guanosine 5'-diphosphate
GMP	Guanosine 5'-monophosphate
GTP	Guanosine 5'-triphosphate
HILIC	Hydrophilic interaction chromatography
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzyme A
HR-MS	High resolution mass spectrometry
IC	Ion chromatography
IMP	Inositol 5'-monophosphate
IP-RP	Ion-pair reversed phase
L. lactis	*Lactococcus lactis*
LC	Liquid chromatography
LC-MS	Liquid chromatography mass spectrometry
M. corallina	*Microbispora corallina*
m/z	Mass-to-charge ratio
Man6P	Mannose 5-phosphate
METLIN	Metabolite Link
MRM	Multiple reaction monitoring
MS/HRMS	High resolution tandem mass spectrometry
MS/MS	Tandem mass spectrometry
NAD^+	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide, reduced
$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced
NIST	National institute for standard and technology
NMR	Nuclear magnetic resonance
OMP	Orotidine 5'-monophosphate
PEP	Phosphoenolpyruvate

ppGpp	Guanosine pentaphosphate
PPP	Pentose phosphate pathway
QqQ	Triple quadrupole
Q-TOF	Quadrupole time of flight
R. mobilis	*Ruegeria mobilis*
R5P	Ribose 5-phosphate
Ribu5P	Ribulose 5-phosphate
RT	Retention time
S. coelicolor	*Streptomyces coelicolor*
S/N	Signal to noise ratio
SIL-IS	Stable isotope-labeled internal standard
S-VG	Vanillin-β-glucoside producing S288C
TBA	Tributylamine
TCA	Tricarboxylic acid cycle
TEA	Triethylamine
TLC	Thin layer chromatography
TOF	Time-of-flight
TP	Time point
UDP	Uridine 5'-diphosphate
UDP-Glc	Uridine 5'-diphosphate glucose
UHPLC	Ultra high performance liquid chromatography
UMP	Uridine 5'-monophosphate
UTP	Uridine 5'-triphosphate
VG	vanillin-β-glucoside
WT	Wild type
XMP	Xanthosine 5'-monophosphate
Xylu5P	Xylulose 5-phosphate
ZMP	5-amino-4-imidazolecarboxamide ribotide
α-KG	Alpha-Ketoglutaric acid

Table of Contents

1. Introduction to the work done during this thesis

The primary aim of this thesis was establishment of LC-MS based analytical methods for qualitative and quantitative analysis of intracellular metabolites from various microorganisms. The establishment process covered all aspects from sample preparation to detection by complementary separation and mass spectrometry detection. The method was subsequently utilized in targeted approaches to quantify the changes in intracellular metabolite concentrations as a response of gene alteration or stress conditions. Furthermore, multitargeted approach was also conducted in order to identify as many intracellular metabolites as possible. The ability to measure the intracellular metabolites can help in understanding the various biological and biochemical processes in the living organisms.

The thesis is divided into two sections: i) introduction and ii) results and discussion. The introduction section describes the importance and the relevant terms in the field of the intracellular metabolites analysis with a focus on the analytical and sample preparation challenges. Resent developments in mass spectrometry and separation techniques used in metabolomics studies are also covered. The results and discussion section covers the results from the analytical and sample preparation method development and the findings from application of the methods for measurement of intracellular metabolites in various microorganisms.

The work in this thesis resulted in 6 peer reviewed and 1 non-peer reviewed papers.

Paper 1 covered the development of the separation and detection method which is applied for the analysis of nucleotides in *Lactococcus lactis.*

Paper 2 applied the developed method for analysis of c-diGMP, a signaling molecule responsible for the shifts between the motile and sessile life in *Ruegeria mobilis.*

Paper 3 applied the method in the analysis of a targeted group of intracellular metabolites involved in the production of vanillin-β-glucoside in *Saccharomyces cerevisiae*.

Paper 4 applied the method for measurement of the central carbon metabolites to investigate the effects of the recombinant protein production on the metabolism.

Paper 5 describes an approach for creating a blank matrix that can be used for spiking and validation of analytical methods for measurement of intracellular metabolites.

Paper 6 covers the multitargeted analysis of the intracellular metabolites extracted from various organisms.

Paper 7 (non-peer reviewed) covers the different steps and challenges in the analysis of intracellular metabolites.

1.1 Introduction to metabolomics

Metabolomics covers the analytical approaches and data evaluation to obtain information about the metabolites in a biological system [1-5]. It can be considered as the comprehensive qualitative and quantitative analysis of all low molecular weight (<1000 Da) metabolites both intracellular and extracellular [1-5]. However due to the physicochemical and structural diversity of the metabolites which include small organic compounds such as organic acids, inorganic ionic species, monosaccharides, amino acids, nucleotides, sugar phosphates etc., this goal has not been reached yet. Despite the primary metabolites, which are important for the life of the cells, a wide range of secondary metabolites exist as well, resulting in approximately 1000 metabolites in *S. cerevisiae* [6,7] or up to 200.000 in plants [2,3,8]. The work in this thesis was focused on the measurement of the primary metabolites that occur inside the cell. Therefore the following section focuses on the importance of intracellular metabolites.

1.1.1 Intracellular metabolites and their importance

In industrial biotechnology, cell factories are used for production of a range of valuable products such as proteins (enzymes or biopharmaceuticals), antibiotics, food additives etc. [9-13]. The production rate and the amounts of the produced products are usually connected with the primary metabolism inside the cells because the primary central carbon metabolism provides precursors, cofactors, energy in the form of adenosine 5'-triphosphate (ATP) and redox equivalents such as nicotinamide adenine dinucleotide phosphate, reduced (NADPH) for the synthesis of the industrially relevant products [9,12]. The central carbon metabolism is a network of reactions that convert the main carbon source into building blocks and energy. According to the classical textbooks, the central carbon metabolism includes the Embden-Meyerhof-Parnas (EMP) pathway of glycolysis, the pentose phosphate pathway (PPP), and

the tricarboxylic acid cycle (TCA) (Figure 1), with individual variations depending on the ecological position in which the organism lives [14-16].

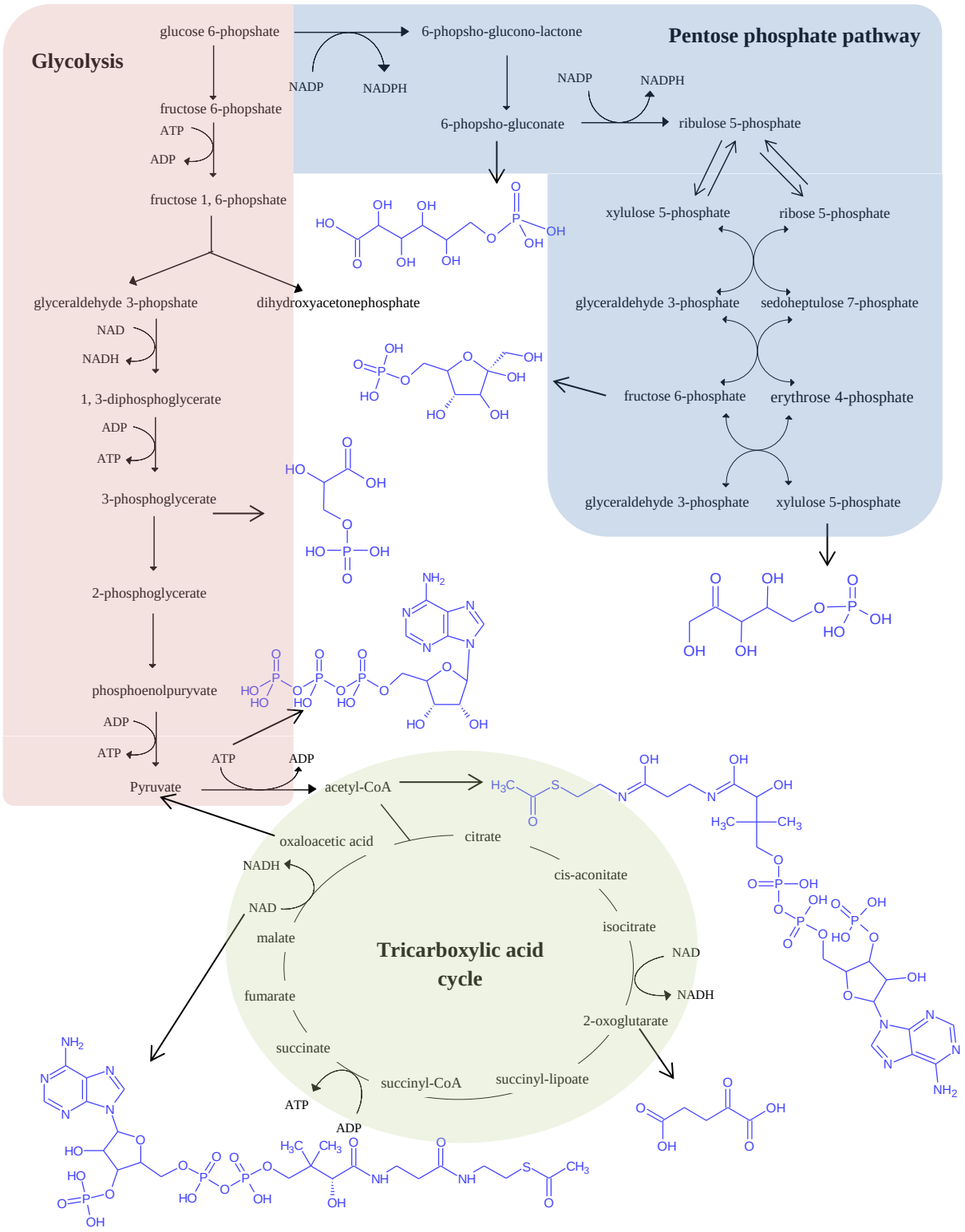


Figure 1. Central carbon metabolism

The relevant intracellular metabolites that are involved in the central carbon metabolism can be classified into:

- carboxylic acids (e.g., pyruvate),
- phosphorylated compounds including sugar phosphates (e.g. glucose 6-phosphate (G6P)),
- phospho-carboxylic acids (e.g., phosphoenolpyruvate (PEP)) and
- nucleotides such as ATP, nicotinamide adenine dinucleotide (NAD^+) and nicotinamide adenine dinucleotide phosphate ($NADP^+$) [17].

Besides being energy sources and building blocks for DNA and RNA, nucleotides can act as signaling molecules as well. Cyclic nucleotides such as cyclic di-3',5'-guanylate (c-diGMP), cyclic di-3',5'-adenylate (c-diAMP), adenosine 3',5'-cyclic phosphate (cAMP) and guanosine 3',5'-cyclic phosphate (cGMP) have key roles in signaling processes in all domains of life [18-21].

Understanding of the primary cellular metabolism can be utilized to increase the production of precursors for the industrially relevant metabolites and for optimization of the biotechnological processes in general. The measurement of the intracellular metabolite pools together with the metabolic flux analysis can help in understanding how the genetic manipulations affect the metabolism of the cell, to identify bottlenecks in target pathways and to predict targets for metabolic engineering [22,23]. Combining the metabolomics data with those from the genomics, transcriptomics and proteomics analysis, helps in understanding the metabolism and the regulations in the cells [13].

1.1.2 Challenges in measuring intracellular metabolites

Three properties of the intracellular metabolites make their analysis challenging [24,25]:

i) The rapid turn-over - due to the fast turn-over rate of the metabolic pools, the intracellular metabolite concentrations can be adjusted rapidly to new levels. Turnover

times of for example the central carbon metabolism are in the range of seconds whereas for amino acids it is in the range of minutes. Therefore, fast sampling and metabolism arrest is necessary for analysis of these metabolites [24-29];

ii) The varied abundance - the levels of the intracellular metabolites in the cell are determined by the function of the enzymes and therefore at a certain metabolic state, some of the metabolites can be present in very high amounts, whereas others will be present in trace amounts [1,24,25];

iii) The physicochemical diversity - the intracellular metabolites exhibit different physicochemical characteristics and can for example contain one or more phosphate or carboxylic groups or both (Figure 1). Furthermore distinguishing between some of the metabolites can also be difficult due to the similarities in their elemental composition and structures [1, 24,25,28,30].

The following sections will cover the approaches and the steps in the analysis of intracellular metabolites.

1.2 Designing a metabolomics experiment-targeted versus untargeted approach

Metabolomics can be mainly divided into two approaches: targeted and untargeted. Which approach will be applied depends on the question that needs to be answered [30-32].

A substantial part of this Ph.D. study was based on the targeted approach where a predetermined set of metabolites related to a specific metabolic pathway of interest were measured. This approach answers the questions concerning the level of a specific analyte in a sample and enables quantification of the metabolites of interest by employing authentic standards for creating calibration curves [32,33]. Consequently this approach is focused on the quantitative changes of the measured metabolites in the cell. The advantages of the targeted metabolomics with QqQ are the increased specificity and sensitivity where metabolites

present in very low amounts can be quantified [31-33]. However it fails in detecting previously uncharacterized compounds because only selected metabolites are measured. There is a wealth of literature where selected intracellular metabolites are targeted and measured in order to address metabolic responses caused by e.g. genetic engineering, modified growth conditions etc. [34-38]. For e.g. the responses of *S. cerevisiae* to the redox perturbations caused by overexpressing nicotinamide adenine dinucleotide, reduced (NADH) oxidase, NADH kinase and transhydrogenase has been investigated using the targeted approach, by measurement of ATP, adenosine 5'-diphosphate (ADP) and the redox cofactors [34]. Targeted metabolomics was also applied in the study conducted by *Boer et al.* [37] where intracellular metabolites were measured in *S. cerevisiae* to investigate the pathways linking the nutrient environment to growth rate. The measurements showed that some of the intracellular metabolites limiting growth include glutamine, ATP, pyruvate and uridine 5'-triphosphate (UTP). Targeted LC-MS metabolomics was also applied for investigating the changes in the metabolite pools due to nutrient depletion in *S. coelicolor*. Decrease of the phosphorylated metabolite pools was observed in the phosphate and glutamine limited cultures, while decline in the amino acid and organic pools was observed in the glutamine limited cultures [38].

Untargeted metabolomics is considered to be a hypothesis free or hypothesis generating approach [31,32]. This approach has the aim of measuring as many known and unknown metabolites as possible, to answer the question what is the metabolic profile of a biological sample [31,32]. High resolution instruments are traditionally applied for untargeted metabolomics [31]. Using these instruments, information on the accurate mass, isotopic pattern and isotope abundances of the compounds in the sample of interest can be obtained. This together with the specific retention time and the accurate mass of the fragments facilitates the identification of the unknowns. The untargeted approach offers the advantage

of less method development compared to targeted approach, but the data generated is more complex and therefore requires additional data analysis [39]. Hundreds of peaks can be detected in a sample thus the manual inspection of the peaks is very time consuming. However with the recent developments within bioinformatics tools, identification of the metabolic peaks has become a relatively automated process. The accurate mass of the metabolite detected by the MS is searched for in a metabolite database such as METLIN [40] (for electrospray ionization tandem mass spectrometry, https://metlin.scripps.edu/index.php) or national institute for standard and technology (NIST) database [41] (for electron impact ionization mass spectrometry, http://www.nist.gov/srd/nist1.htm). The database match is considered as a putative identification of the metabolites. Thus the result must be confirmed by comparing the fragmentation pattern of the compound that has been assigned to the peak, to the measured fragmentation pattern for the particular compound in the sample of interest [31,40]. Untargeted metabolomics has been used to study the changes in the intracellular metabolic profile of the human liver cell line HEPG2 as a response to the exposure to a selected toxicant [42].

Despite of the approach chosen and the question that needs to be answered the analysis of the intracellular metabolites can be divided into three steps: i) sample preparation, ii) sample analysis and iii) data analysis and interpretation. During method development (which was the main part of this thesis) the work flow usually starts with optimization of the metabolite detection followed by establishment of the sample preparation method. Therefore techniques for analysis of the intracellular metabolites will be discussed first followed by the most commonly used sample preparation methods.

1.3 Analytical platforms used in metabolomics

Mass spectrometry in combination with a separation method is the key technology for analysis of intracellular metabolites. The recent significant progress in MS based

metabolomics gives the researchers opportunities to choose between different separation techniques such as capillary electrophoresis (CE), gas chromatography (GC) and liquid chromatography (LC) [1,4,43-45]. MS in combination with a sample preparation and separation technique has high sensitivity and wide linear dynamic range and is therefore often used for analysis of intracellular metabolites.

Nuclear magnetic resonance (NMR) is another technology applied for metabolomics studies. The advantage of NMR is the minimal requirement for sample preparation, it is not destructive, it is useful in structural characterization of unknowns but it has low sensitivity [1, [46,47].

1.3.1 Separation techniques

The use of a separation technique is essential for metabolomics studies due to the large number of metabolites to be analyzed and the complexity of the biological samples. The choice of separation technique (GC, CE or LC) to be applied depends on the initial goal of the study and the metabolite class of interest. Table 1 lists some of the advantages and disadvantages of the most commonly used separation techniques in combination with MS.

Table 1. Advantages and disadvantages of the most commonly used mass spectrometry based analytical techniques in metabolomics [1,17,48-50].

Technique	Advantages	Disadvantages	Metabolites analyzed
GC-MS	• Suitable for volatile and semi volatile analytes • Good and reproducible separation • Databases available due to the reproducibility of the fragmentation	• Not suitable for non-volatile and thermally labile compounds • Requires derivatization which is time consuming • Derivatization can cause difficulties in the identification of unknowns	Amino acids Carboxylic acids Purines and pyrimidines Sugar phosphates
CE-MS	• Good separation • Requires small sample volumes • No buffer gradient is applied thus no fluctuation in the electrospray ionization	• Interfacing difficulties with MS • Incompatible buffers with MS • Poor sensitivity • Migration time shifts	Organic acids Sugar phosphates Nucleotides Coenzymes
LC-MS	• Derivatization not necessary • Suitable for polar, semi polar and non-polar metabolites • Offers a wide range of stationary phases with different functionalities (RP, hydrophilic interaction chromatography, ion-exchange) • Allows analysis of thermally labile analytes	• Electrospray ionization suffers from matrix effects • Restrictions on eluents (only volatile buffers and additives can be used) • Fragmentation not reproducible as in GC-MS	Organic acids Sugar phosphates Nucleotides Coenzymes

GC is widely used in metabolomics studies due to the high separation efficiency and the easy interfacing with MS [1,50,51]. GC is primarily used for thermally stable and volatile metabolites. Chemical derivatization (e.g. silylation) is necessary for semi-volatile compounds, which is considered as a time consuming step that complicates the sample preparation and increases the variances in the analysis [48]. In addition, a considerable number of very polar metabolites cannot be analyzed due to the non-volatility. Simultaneous analysis of amino and organic acids using GC-MS was shown to be possible using

chloroformate derivatization [50], while for sugar phosphates using silylation [51]. *Sellick et al.* [52] described a protocol for analysis of several intracellular metabolites using GC-MS, however it was not possible to analyze the thermally labile metabolites such as ATP, ADP, NAD^+, nicotinamide adenine dinucleotide, reduced (NADH).

CE provides efficient separation of charged metabolites and has the advantage of a small volume required for analysis. It is not often used in metabolomics analysis but it is becoming a promising technique as shown by several published CE-MS methods for analysis of intracellular metabolites [30,53,54].

Büsher et al. [44] made an extensive comparison of GC-MS, CE-MS and LC-MS for the analysis of 75 intracellular metabolites such as sugar phosphates, nucleotides, coenzymes, redox cofactors, amino and organic acids. The three separation techniques were compared in terms of metabolite coverage, matrix effects, separation of isomers, sensitivity, analysis time and reproducibility, with LC giving the best compromise between the tested parameters.

The coupling of LC with MS was facilitated with the introduction of atmospheric pressure ionization techniques such as electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI). The coupling of LC with MS requires compatible eluents. Non-volatile buffers and additives (phosphate, Na^+, K^+ etc.) should not be used as it can affect the evaporation thus reducing the signal intensity. Reversed phase chromatography uses solvents compatible with MS (volatile buffers e.g. ammonium acetate, organic solvents e.g. methanol, acetonitrile). The combination of RP and acidic mobile phases has been shown to work for a number of nitrogen containing compounds such as amino acids, purine and pyrimidine bases and monophosphorylated nucleotides, but not for di and triphosphorylated compounds [55]. The alternatives to RP for separation of intracellular metabolites are: hydrophilic interaction chromatography (HILIC), AEC, IC and IP-RP. Each of these separation modes has its own advantages and disadvantages and some of them are summarized in Table 2.

Table 2. Advantages and disadvantages of the most commonly applied LC modes for separation of intracellular metabolites [17,56-59]

	Advantages	**Disadvantages**
HILIC	• enhanced sensitivity with MS detection (compared to RP)	• cannot distinguish between structural isomers • reproducibility issues • poor peak shape
AEC	• can retain anionic compounds	• high content of salt in the eluent not compatible with ESI • difficulties in elution of well retained compounds with multiple charges
IC	• can retain anionic compounds • good reproducibility	• some metabolites are not stable at the high pH required for elution • MS not compatible solvents-suppressor needed
Ion pair	• can retain anionic compound • can separate structural isomers • more reproducible and better peak shape than HILIC • MS compatible eluents	• system contamination • long equilibration time • operation in only one MS mode (positive or negative)

IP-RP was used during this Ph.D. study for separation of the intracellular metabolites of interest and is therefore discussed in more detail. IP-RP is an alternative to ion-exchange chromatography and is usually considered as a modification of RP chromatography for retention of ionized compounds. IP-RP has been shown to be more reproducible, to give better separation especially for structural isomers and better peak shape than HILIC [17,57]. Hydrophobic stationary phases and solvents containing ion-pair reagents are used to separate the ionic compounds in IP-RP. For retention of negatively charged compounds, alkylamines are typically used. The non-volatile tetraalkylammonium salts that were used in the past as ion-pair reagents, have nowadays been replaced by more volatile reagents such as dibutylamine (DBA) and tributylamine (TBA) [17,57,60-63]. The alkyl chains of the ion-pair reagent interact with the hydrophobic functionalities from the stationary phase while the amino group interacts with the negatively charged analyte. One of the hypotheses for IP-RP

retention assumes that the formation of the ion-pairs happens in the solution, which then partitions between the mobile and the stationary phase (Figure 2A) [60].

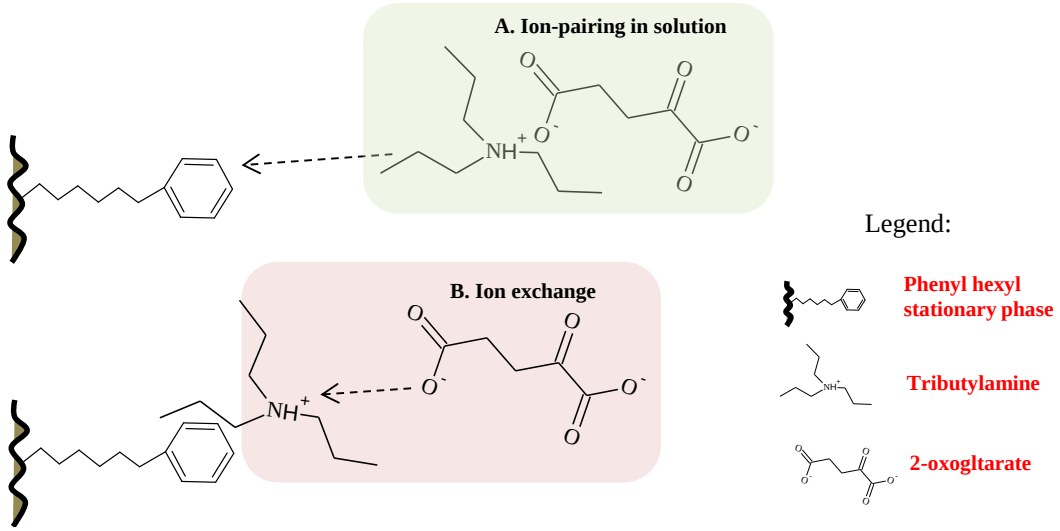


Figure 2. Schematic representation of the two mechanisms of retention in ion-pair chromatography.

The second proposed mechanism, that is most commonly accepted, assumes that the hydrophobic part of the ion pair reagent adsorbs on the hydrophobic stationary phase to form a dynamic ion-exchange surface. The analyte is then retained on this surface by ion-exchange mechanism and formation of ion-pairs between the ionic compounds and the ion pair reagents with opposite charge as shown in Figure 2B [60].

IP-RP has been widely applied for the analysis of intracellular metabolites. *Luo et al*. [17] reported simultaneous analysis of the central carbon metabolites from *Escherichia coli* (*E. coli*) using IP-RP-LC tandem mass spectrometry (MS/MS). This paper also covers some of the most important method development steps in IP-RP such as type and concentration of the ion-pair reagent, the pH of the mobile phase, as well as type and strength of the organic solvent. The analysis of nucleotides in yeast using ion-pairing for separation has also been previously reported by *Seifar et al.* [61], while *Coulier et al.* uses the IP-RP separation method for analysis of the same metabolites in bacteria [63].

Despite the advantages of IP-RP over the other separation techniques for intracellular metabolites, several issues need to be taken into account before choosing this type of chromatographic separation. This technique can retain and separate only charged molecules. The ion pair reagent can never be fully washed from the column and the LC system, thus one should dedicate a particular column and LC to ion-pair application. Furthermore, since the ion-pair reagent can cause ion suppression, the MS needs to be operated in an ionization mode opposite to the charge of the ion-pair reagent used in the mobile phase.

1.3.2 Mass spectrometry detection

Various types of mass spectrometers are nowadays used in metabolomics approaches. They generally differ in the ionization source and the mass analyzer. The commonly used ion sources are electron ionization (EI), APCI and ESI [45,47]. EI is used in connection with GC and is a so called ‘’hard’’ ionization since it causes fragmentation of the molecular ion. The fragmentations are characteristic and reproducible, which allows identification of the MS peaks by comparison to those in public and commercial databases (e.g NIST that contain generalized collection of chemicals or FiehnLib for metabolites) [41,64]. ESI (for polar molecules) and APCI (for less polar molecules) are considered as ‘’soft’’ ionization techniques due to the low energy to which the analytes are exposed during the ionization thus resulting in less fragmentation [45,65]. ESI works well with polar molecules and is thus well suited for the intracellular metabolites. In ESI the liquid sample is nebulized, evaporated by applying heat and nitrogen gas and ionized under atmospheric pressure in a strong electric field (Figure 3) [65].

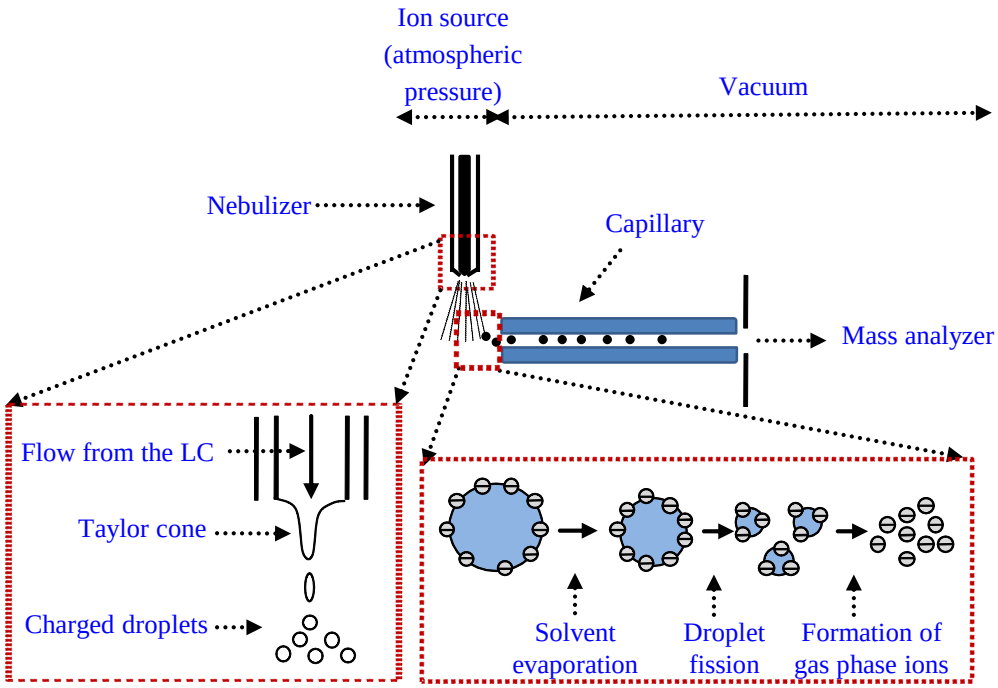


Figure 3. Scheme of electrospray ionization.

A well-known problem with ESI is the matrix effect (ion suppression or enchancement) caused by interfering compounds that elute at the same time as the metabolite of interest [66]. The matrix effects are usually corrected by using SIL-IS [61,66,67] which will be discussed in more detail in the method validation part.

The mass analyzer provides separation of the ions according to their mass-to-charge ratio (m/z). The most important criteria for the mass analyzers are: sensitivity, mass resolution, accuracy, scan speed or acquisition rate and MS/MS capabilities [65]. The most commonly applied analyzers in metabolomics studies are the quadrupole and the time-of-flight (TOF). The quadrupole mass analyzers are robust, relatively cheap and easy to use. It can act as a filter and let only ions with a certain m/z pass or it can act as a scanning instrument where ions of different m/z are detected consecutively thereby a mass spectrum is obtained (full scan mode) [65]. The operation principle in TOF involves measuring the time required for an ion to travel down a flight tube to the detector. TOF provides high resolution and mass

accuracy which allows reliable assignment of the measured masses to the elemental composition of a compound [45,65].

The mass analyzers can be combined to a tandem mass spectrometer (MS/MS), where two mass analyzers are combined with a collision cell in between. The tandem mass spectrometers used during this Ph.D. were the QqQ and the Q-TOF MS. Instrumentally the QqQ contains two quadrupoles mass analyzers arranged in series, with a collision cell in between which can be a quadrupole or a hexapole (Figure 4). QqQ offers high specificity, sensitivity and high dynamic range [68].

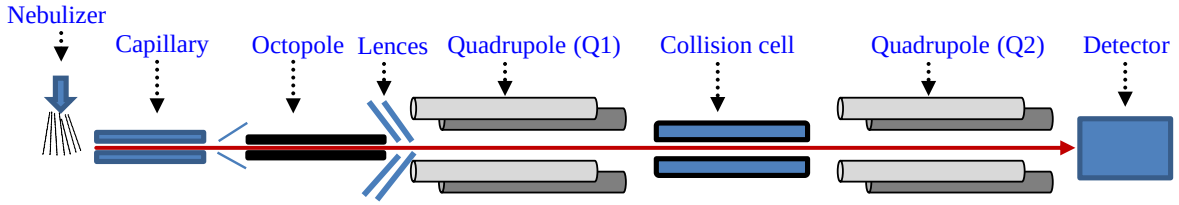


Figure 4. Schematic overview of a QqQ mass spectrometer. Figure modified from Agilent 6400 Series Triple Quadrupole LC/MS, Concepts Guide [69].

When compared to the QqQ, in the Q-TOF instrument the third quadrupole has been changed by TOF mass analyzer (Figure 5). Q-TOF can be operated as a TOF or as a tandem mass spectrometer where a precursor ion is isolated by a quadrupole mass analyzer, fragmented in the collision cell and the fragment spectrum is acquired in the TOF mass analyzer. A Q-TOF instrument combines the high mass resolution of a TOF instrument with the MS/MS capability facilitating the tentative identification of unknown compounds.

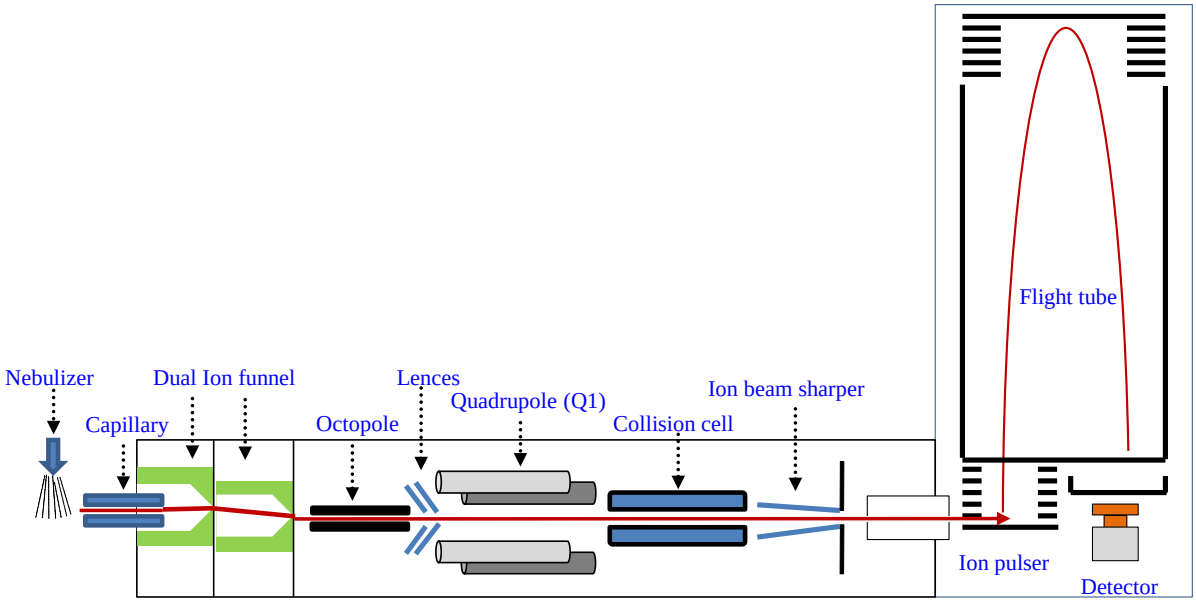


Figure 5. Schematic overview of a QTOF-MS instrument. Figure modified from 6200Series TOF and 6500 Series Q-TOF LC/MS System, Concepts Guide [70].

1.3.2.1 Multiple reactions monitoring versus full scan high resolution MS

QqQ MS are traditionally applied in the targeted approach and are typically operated in MRM mode (Figure 6A) [43]. In development of an MRM method a precursor ion, product ion and collision energy are optimized for each analyte to give the best signal. Careful selection of the product ions is necessary especially for structural isomers as these compounds often produce very similar product ions. In an MRM mode the first quadrupole selects the precursor ion of interest, the second quadrupole fragments the precursor ion, while the third quadrupole isolates the proper product ion (Figure 6A).

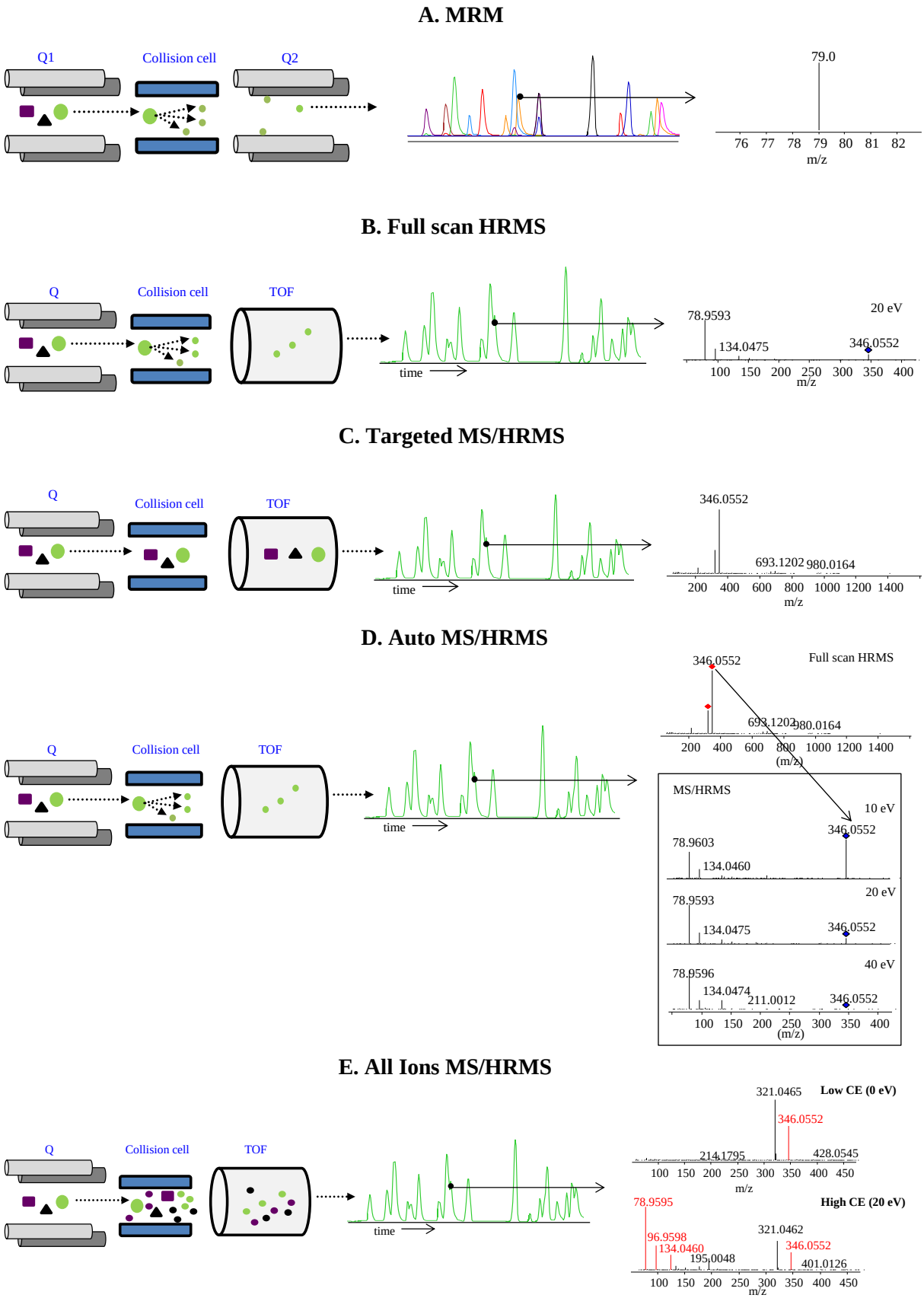


Figure 6. Schematic overview of the different acquisition modes used during this Ph.D. showing the difference in the data obtained

This process is repeated for each metabolite in a cyclic manner and has the advantage to determine selected metabolites in a few minutes from a small amount of sample. One of the disadvantages of the MRM is that the number of metabolites that can be measured is limited. By introducing too many MRM transitions into the method, the scan time per MRM needs to be lowered. This will compromise the signal to noise (S/N) ratio and the number of data points across the peak (ideally 10-20) important for reliable quantification (Figure 7) [43].

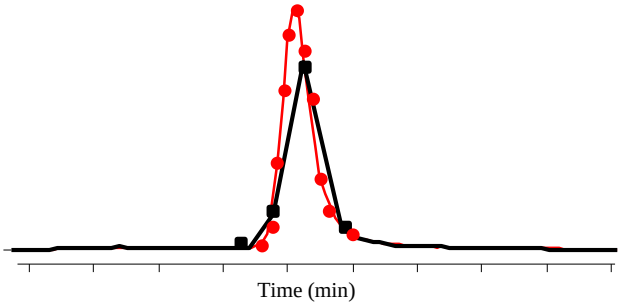


Figure 7. Overlaid chromatograms showing 11 data points (red color) and 3 data points (black color) across a chromatographic peak.

The so called dynamic MRM (DMRM) acquisition mode of the Agilent QqQ (called *Scheduled* MRM^{TM} in the QqQ from AB Sciex) offers an increased number of metabolites monitored per run. In DMRM or *Scheduled* MRM^{TM} mode, the chromatographic run is divided into time segments where the MRM transitions of a specific compound are monitored only in the time segment where the particular compound elutes. The advantage of this operating mode is lower number of concurrent transitions, longer dwell times per particular MRM, better peak symmetry, increased sensitivity and better S/N.

Another approach to increase the number of metabolites monitored is the full scan MS experiments, where all ions generated from the sample are measured (Figure 6B). However, full scan measurements on a low resolution and low scan speed instrument such as the QqQ instruments are not optimal, especially for low mass compounds, due to interference from contaminants with the same nominal mass [43]. Furthermore, the signal of the low abundant

compounds will be hidden by the high background noise compromising the quantification. Full scan measurements on a high resolution instrument such as Q-TOF MS (or Furrier Transform MS such as Orbitrap) overcome these issues due to the high mass accuracy and the mass based separation.

Q-TOF offers both full scan HR-MS (Figure 6B) and MS/HRMS measurements (Figure 6C, D and E). The HRMS measurements are conducted when accurate mass MS data or determination of precursor ion masses for subsequent MS/MS are needed. The MS/MS data can be obtained by the targeted MS/MS, auto MS/MS or all ions MS/MS (MS^E in Waters Q-TOF instruments).

Targeted MS/MS involves manual selection of the precursor ion based on the MS data and is used when the precursor ion of interest is known (Figure 6C). Using targeted approach better selectivity can be achieved and by comparing the fragmentation pattern with those in a database, identification of the compound can be achieved [71,72].

Auto MS/MS approach is used for obtaining MS/MS data of complex samples where the operator does not know which precursor ions to choose. The auto MS/MS mode provides MS and MS/MS data from a single run, by combining MS scan cycles with MS/MS scans of selected precursor ions depending on their abundance in the MS scan (Figure 6D) [65]. The disadvantage of this mode is that not all metabolites detected will be fragmented and the more concentrated metabolites will be preferred for fragmentation than the less concentrated. However, exclusion or inclusion lists for selection of precursor ions can help in increasing the fraction of metabolites that can be fragmented thus providing useful MS/MS data.

The advantage of all ions MS/MS approach is that all the ions despite of their intensity will be fragmented which eliminates the need of specifying the precursor ions (Figure 6E) [64]. The data are acquired at both low and high collision energy. The low value produces the precursor ions for the compounds and the high value generates the precursor and their

product ions. The fragmentation patters can be used to reveal structural information about the known and unknowns. However this approach was not very suitable for the analysis of intracellular metabolites since many of them produce similar fragments which complicate the identification of the metabolites.

During this Ph.D. study MRM (**Paper 1**-**5**), HRMS or MS/HRMS (**Paper 6**) measurements were used. MRM was used for the quantitative measurement of the selected list of metabolites relevant for the particular study. The high resolution measurements were used for detection and putative identification of as many metabolites as possible in the biological extracts.

1.4 Sample preparation

Due to differences in cell structure the sample preparation is organism dependent and it is therefore difficult to establish a general sample preparation method for both prokaryotes and eukaryotes [25]. During this thesis I have worked with organisms from both domains thus an overview of the sample preparation techniques for both prokaryote and eukaryote will be given. The sample preparation for intracellular metabolites analysis consists of four steps: rapid sampling, quenching, extraction and sample concentration. The rapid sampling is usually achieved by in-house developed devices and not commercially available ones. The main characteristics of these devices are the sampling time and the reproducibility of the sampling [28,73].

1.4.1 Quenching

After sampling, the metabolism of the cells needs to be inactivated (quenched) in order to get a representative sample of the physiological state of interest. The quenching is usually evaluated by the adenylate energy charge ratio (ECR) given by the formula $(\mathrm{ATP} + 0.5\mathrm{ADP})/(\mathrm{ATP} + \mathrm{ADP} + \mathrm{AMP})$ [74,75]. For a growing cell the ECR is in the range

of 0.80-0.95 and indicates high amount of ATP relative to ADP and AMP, thus a physiologically healthy cell [75]. In general there are two quenching strategies [28]. The first strategy allows separation of the quenching supernatant from the biomass, whereas the second one does not allow that separation, thus the quenching and extraction are combined. In both cases the quenching is usually done by changes in the pH (below 2 or above 10) or temperature (-40°C or 80°C) [25]. During this study, quenching using cold MeOH /water, cold MeOH/glycerol, formic acid, cold buffered and non-buffered saline (0.9% (w/v) NaCl) were tested.

In the literature there is no agreement on which quenching technique gives the best balance between the main problems that arise during the quenching such as physical or chemical alterations of the metabolites, metabolite leakage and contamination. Table 3 lists some of the quenching procedures found in the literature, together with their advantages and disadvantages.

Table 3. Advantages, disadvantages and applications of commonly used quenching procedures.

Quenching	Organism	Advantages	Disadvantages	References
Acid/bases	• Bacteria • Yeast	• Combined quenching and extraction therefore no losses due to leakage	• Media components can interfere with the analysis • Extracellular and intracellular metabolites are not separated	[76-78]
Cold MeOH	• Bacteria • Yeast • Filamentous fungi	• Reproducible and simple • Allows separation of the intracellular metabolites from the extracellular	• Leakage especially when applied to bacteria	[35-37,61,79-83]
Glycerol (combined with methanol or saline)	• Bacteria • Yeast	• Cryoprotectant-no harm to cells and membranes • Lower temperatures than -40 °C can be obtained when combined with MeOH	• High boiling point, difficult to remove it from the quenching supernatant-leakage cannot be evaluated • Residues of glycerol result in viscous extracts difficult to inject in LC-MS • Requires washing step for removing the glycerol residues	[84,85]
Cold saline	• Mammalian cells	• The temperature applied (0 °C) are mild • Combined with filtration suitable for amino acids analysis	• It is not suitable for metabolites with fast turn over (e.g. phosphorylated)	[24, 86]

One important requirement for the quenching approach is to avoid losses of the metabolites due to leakage or residual metabolic activity. The chemical alteration of the metabolites during the sample preparation is also an issue and it is not always taken into consideration during the method development. *Ortmayr et al.* [87] and *Sporty et al.* [88] showed that the sample preparation procedure is the major contributor to the overall measurement uncertainty of the redox cofactors. Therefore detailed study on the stability of the analytes during the sample preparation is necessary for better accuracy.

Yeast has been considered to be more stable during the quenching when compared to the bacteria. The first attempt at separation of the quenching supernatant from the yeast cells was done by *Sáez* and *Lagunas* [89] by using fast filtration followed by washing the cells with -40 °C MeOH/H_2O (60/40, v/v). Later *De Konin* and *Van Dam* [90] improved this technique by sampling yeast directly into -40 °C MeOH/H_2O (60/40, v/v) followed by centrifugation to separate the cells. Nowadays this quenching technique is one of the most frequently used methods for sub second arrest of the enzymatic activities in the cells. However it has been shown that yeast is also prone to leakage during the conventional cold MeOH/H_2O quenching which can be reduced by using pure MeOH at -40 °C or lower [81]. Furthermore, decreasing the exposure time of the yeast cells to the cold MeOH/H_2O quenching solution will decrease the loss of the intracellular metabolites as a result of leakage during the quenching. Several other alternatives to the cold methanol technique have also been suggested in the literature. *Villas-Bôas et al.* [84] showed that by using cold glycerol/saline solution for quenching of yeast and bacteria better metabolite recoveries can be obtained when compared to the MeOH/H_2O. *Link et al.* [85] showed that the combination of glycerol and methanol is a better option due to the lower viscosity of the quenching solution as well as possibilities for sample handling at temperatures lower than -40 °C. The glycerol/MeOH combination showed reduced leakage especially for ATP in *E. coli* when compared to the MeOH/H_2O quenching.

Today bacteria and filamentous fungi are known to be more susceptible to leakage when using the current quenching procedures than for example yeast and thus the quenching of bacteria is a challenging task [83,91]. If possible cell separation from the quenching solution is usually avoided for bacteria. Thus the measurement of the intracellular metabolites requires preparation of two different samples, one from the entire culture and one from the culture supernatant. The intracellular metabolites are then determined by subtraction of the levels of extracellular metabolites from the sum of intra and extracellular metabolites [82]. However this has been shown to suffer from large standard deviations. More recent example of a quenching method based on pH change, where the quenching and the extraction were combined, was reported by *Jendersen et al.* for *L. lactis* using cold 10 M formic acid [92]. The quenching showed to be suitable for analysis of nucleotide phosphates.

Due to the low amount of the intracellular metabolites and high amount of the extracellular compounds e.g. media components, the separation of the cells from the quenching supernatant is a critical issue. In general the separation of the biomass from the quenching supernatant is achieved by either centrifugation or filtration [24]. In this respect, the time used for separation is the crucial parameter. During the centrifugation the cells are exposed to low temperatures and a cold shock phenomenon might induce leak of the metabolites in the quenching supernatant [91]. Fast filtration of the bacterial cells without quenching followed by wash using saline solution have been shown to be a reliable quenching method for analysis of amino acid pools [91]. However, the washing step that is used to remove the residues from the extracellular metabolites can also induce leakage and loss of the intracellular metabolites [93,94]. Some studies showed that filtration is not a fast enough quenching method for analysis of the metabolites with fast turn-over rate such as the phosphorylated metabolites [91]. Furthermore blockage of the filter, limits the amount of biomass that can be separated from the quenching supernatant [24]. Care must also be taken

regarding the pore size of the filters that might result in loss of the cells into the filtrate. In respect to the quenching of filamentous fungi, *Jonge et al.* [95] showed that cold MeOH/H_2O used for arresting the yeast metabolism, can also be applied for quenching of *Penicillium chrysogenum*. In this study decreasing the MeOH percentage from 60% to 40% (v/v), consequently the temperature from -40 to -25°C, resulted in reduced leakage.

Mammalian cells have been shown to be more fragile than bacteria and yeast due to the lack of a cell wall [24]. Thus, a rapid quenching method for animal cells should retain cell integrity and be compatible with the separation steps used to remove culture medium from the cells. The conventional cold MeOH/H_2O method has been shown to damage the cell membrane of the mammalian cells resulting in leakage and loss of the intracellular metabolites in the quenching medium [24]. *Dietmar et al.* [24] tested several different quenching solutions for mammalian cells such as buffered and non-buffered -40 °C MeOH/H_2O (60/40, v/v) and 0.9 % (w/v) NaCl (0 °C), with the latter showing the least damage of the cell membrane and lowest leakage of the intracellular metabolites into the quenching solution. Furthermore, the centrifugation has been shown to be more suitable for the cell separation from the quenching supernatant than the filtration. The reason for the reduced efficiency of the filtration when compared to centrifugation was speculated to be the three-way interaction among cells, filter, and the quenching solution.

Other quenching procedures given in the literature use cold EtOH/H_2O mixture or liquid nitrogen as quenching solutions [96,97].

1.4.2 Extraction

The step following the metabolism arrest is the extraction of the intracellular metabolites that aims to disrupt the cell structure and release the metabolites from the interior of the cell. The extraction method should extract the metabolites in their original state in a quantitative manner and should prevent any chemical or physical alterations. As for the quenching there is

no standardized extraction method and a range of different solutions and procedures have been described in the literature (for references see Table 4). Furthermore, due to the high chemical and physical diversity of the intracellular metabolites, each solution and condition for extraction may favor a limited range of metabolites and work best with a certain cell type. Combinations of different organic solvents, elevated temperature in combination with boiling solvent (H_2O, EtOH or MeOH), acidic or basic solutions and freeze-thaw cycles are some of the extraction methods for intracellular metabolites (for references see Table 4). Table 4 summarizes the advantages and disadvantages of the commonly used procedures their advantages and disadvantages.

Table 4. Advantages and disadvantages of the most commonly used extraction procedures as well as their applications

Extraction	Organism	Advantages	Disadvantages	References
Boiling EtOH	• Yeast • Bacteria • Filamentous fungi	• Easy and fast procedure • High temperature increases efficiency and denaturation of enzymes • Ethanol is not toxic	• Some metabolites are not stable at high temperatures	[35-37,62,81]
MeOH/chloroform	• Yeast • Bacteria	• Extraction of both polar and non-polar metabolites • Denaturation of proteins by chloroform • Extraction at low temperature - suitable for thermally labile compounds	• Labor intensive • Chloroform is toxic	[27,34]
Acid/bases	• Bacteria	• Usually combines quenching and extraction therefore no losses due to leakage	• Some metabolites are not stable under extreme pH • Neutralization step is required for e.g. perchloric acid and can result in reduced recovery • Freeze/thaw necessary to increase the efficiency of the extraction	[77,98-100]
Cold pure methanol	• Bacteria	• Easily removable by evaporation • Extraction at low temperature suitable for thermally labile compounds	• Low recoveries	[97]

Combination of different polar and non-polar organic solvents is frequently used for extraction of intracellular metabolites. The combination of buffered methanol-water mixture together with chloroform at low temperatures has been used for extraction of intracellular metabolites from yeast and bacteria [27,34,101]. The advantage of this extraction is that it allows extraction of two big groups of metabolites, polar and non-polar, under mild conditions at low temperature but it has been considered as tedious and time consuming. Although some studies report that this combination of solvents result in poor recovery of some of the metabolites [102], others find this method to be optimal and comparable with the other extraction methods [24,101]. Boiling ethanol is a very popular extraction method especially for yeast [35-37,61,81]. The high temperature increases the extraction efficiency of ethanol as well as the denaturation of the enzymes. Furthermore ethanol is less toxic than chloroform. However some studies have reported poor recoveries of the phosphorylated metabolites and tricarboxylic acids due to high temperatures applied during extraction [24,27]. On the contrary, *Canelas et al.* [101] reported that the metabolite recoveries were similar when boiling ethanol and MeOH/chloroform were used for extraction of the intracellular metabolites from yeast. Acids (perchloric and hydrochloric acid) or alkali solutions (potassium and sodium hydroxide) have also been used for extraction of the metabolites that are stable at extreme high or low pH. Losses of many primary metabolites have been demonstrated by using the extraction at extreme pH [27,99]. Furthermore the acid-base methods due to salt production during neutralization are problematic for mass spectrometry applications [27,99,101]. To increase the permeability of the cells, freeze-thawing is often included in the extraction processes and has also been combined with the approaches where the cells are not separated from the quenching supernatant prior to extraction [76,92]. Mechanical disruption of the cell with a bead beater was used by *Sporty et al.* [88] for

analysis of NAD^+ and NADH from yeast in combination with ice cold nitrogen saturated ammonium acetate to reduce the oxidation of NADH to NAD^+.

However, it should be noted that it is practically impossible to avoid losses of the metabolites during the extraction. Furthermore the diversity of the analytes makes the simultaneous extraction and determination of all intracellular metabolites impossible. Therefore the extraction method applied should be a compromise between the reproducibility, metabolite recoveries and compatibility with the analytical method.

1.4.3 Concentration of the extract

The most commonly used approaches for sample concentration is freeze drying (lyophilization), vacuum and nitrogen evaporation. Freeze-drying is commonly used for aqueous samples and has the advantages that the solvent is evaporated without using heat which reduces the possibility of heat degradation of the metabolites. However the freeze-drying can be time consuming and is not very suitable if organic solvents have been used during the extraction [25].

An alternative to the freeze-drying is the vacuum or nitrogen evaporation. This type of evaporation is more suitable for organic solvents than aqueous samples. The evaporation of the aqueous samples can take a longer time and usually requires heating which is not suitable for thermally labile compounds.

For sample purification solid phase extraction (SPE, DSPE) is commonly applied especially in the case of complex matrices such as biofluids, tissue samples and environmental samples (lake sediment) where the quantification can be hampered by matrix effects [103-105]. Anion exchange SPE has also been shown to work well for phosphate removal, which can interfere with the analysis of the intracellular metabolites causing ion suppression [105]. Charcoal has been used in dispersive solid extraction to separate the aromatic from the non-aromatic intracellular metabolites [78,92].

1.5 Analytical method validation and quantification

The main objective of any analytical measurement is to obtain reliable and reproducible data avoiding false positive or negative results. The validation of the analytical methods provides information on the quality, reliability and reproducibility of the established analytical method. When using LC-MS, the general validation approach involves determination of: i) linearity, ii) precision and accuracy/recovery, iii) limit of detection, iv) matrix effects [106-109].

Linearity. The linearity of the analytical method is investigated by creating calibration curves in at least three orders of magnitude. In general three approaches for creating the calibration curves were used during this Ph.D.: i) external calibration, ii) the standard addition approach and iii) calibration using internal standards (Figure 8) [110].

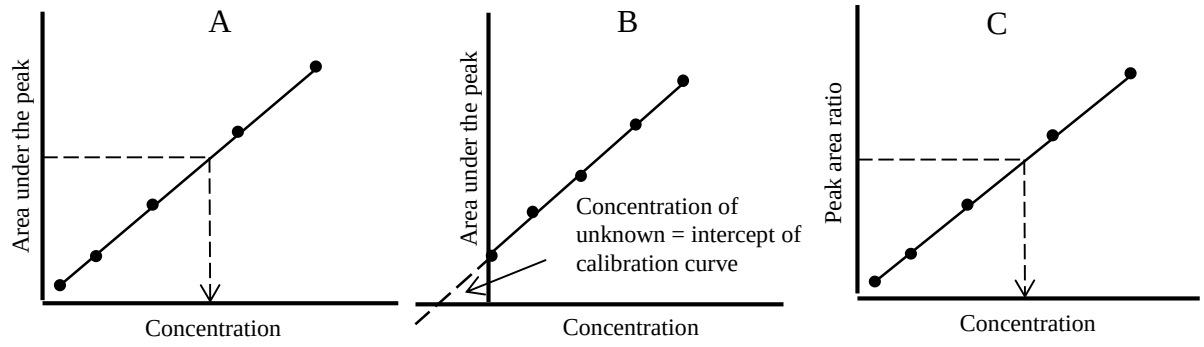


Figure 8. Schematic illustration of the calibrations using the A. external calibration; B. standard addition and C. calibration using internal standard

In the extracellular calibration approach, the calibration curve is prepared in a surrogate matrix. The surrogate matrix can be a neat solution, growth media, extract from a mutant etc. that does not contain the compound or contains traces of it. In this approach a stock solution of the compound of interest is prepared in the surrogate matrix which is then diluted to prepare the calibrants.

Luo et al. [17] and *Bennette et al.* [57] used the standard addition method to quantify metabolites from the central carbon metabolism in *E. coli* and Synechococcus sp., respectively, where known increasing concentrations of analytes were added to individual aliquots of the sample of interest. The concentration of the analyte in the sample was determined from the intercept of the calibration curve (Figure 8B). However, the standard addition approach can be quite laborious, time consuming and require extensive amounts of samples, especially when multiple samples need to be quantified. Furthermore, for compounds that are present in high amounts in the sample such as ATP, saturation of the detector might occur at the high concentration end when using the standard addition method.

The approach using internal standards usually involves addition of known amounts of SIL-IS to each of the calibration solutions and to the samples to be quantified. The quantification is then based on the construction of a calibration curve by plotting the ratio of the peak area of the unlabeled and labeled analyte versus the concentration [61].

Precision and accuracy/recovery. The precision and accuracy/recovery can be assessed using spiked and unspiked samples that are processed through the sample preparation that are quantified against a calibration curve. The unspiked samples in this case are used to check for any presence of the analyte in the spiking matrix. Usually the precision and the accuracy/ recovery are determined on three different levels in triplicates and on three different days in order to check for inter- and intra-batch variability.

Limit of detection (LOD). LOD is the smallest amount of analyte in the sample that can be detected. There are several approaches for determination of the LOD such as using the S/N ratio, from the standard deviation of the blank or from the calibration curve [111].

Matrix effects. As previously mentioned when ESI is used for ionization, suppression or enchancement (matrix effects) of the analyte signal might be observed due to coeluting compounds present in the biological sample [66,76,112]. When using external calibration, due

to differences in the matrix the obtained MS signals will be different for the same compound in the calibrants and the samples. This can cause over or underestimation of the concentration. The suppression or enhancement of the signal is especially observed when quenched whole broth samples of leaky cells are analyzed and even more prominent when rich media is used for cultivations. Therefore the validation and quantification methods for intracellular metabolites are more complicated and less straightforward since it is difficult if not impossible to find a true biological matrix that will be free of the analytes of interest.

The standard addition method corrects for possible matrix effects because it uses the authentic matrix to prepare the calibrants. The method using internal standards is also capable of correcting for matrix effects. Both the standard and its SIL-IS will be equally affected by the matrix effects resulting in unchanged peak ratio that is plotted against the concentration to create the calibration curve.

The suppression/enhancement of the signal depends on the chemical structure of the molecule thus SIL-IS are used due to their structural and physico-chemical similarity with the analyte of interest. However, SIL-IS are not always available and they can be very expensive therefore structural analogues are sometimes used as internal standards [112]. *Stokvis et al*. [67] compared the use of structural analogues and SIL-IS for several anticancer agents using ESI-MS and MS/MS techniques. SIL-IS showed to be preferred over the structural analogue for accurate quantification. In some of the case studies, performance of the assay improved significantly after substitution of a well-functioning analogous internal standard with a SIL-IS. Due to the unavailability of SIL-IS for many of the intracellular metabolites, *Mashego et al.* [113] proposed an alternative way of obtaining SIL by cultivating microorganisms on U-^{13}C labeled substrates and subsequent extraction of the metabolites. The U-^{13}C labeled extracts can serve as SIL-IS and can be added before the extraction to the unlabeled cell samples to correct for any possible losses during the sample preparation and for matrix

effects. However it should be noted that some of the labeled metabolites can be present in very small amounts in the cell extracts and this can create difficulties related to their detection and their use as SIL-IS.

1.6 Data analysis

The workflow used for analysis of the data depends on the question that needs to be answered and the type of the data acquired. The metabolomics studies usually aim for assessing or discovering important differences between groups of samples. The analytical technologies used in metabolomics such as LC-MS produce large amounts of data where statistical and computational methods are used to evaluate this data. This typically includes univariate (ANOVA, t-tests etc.) or multivariate statistical approaches (principle component analysis (PCA), clustering, partial least square regression (PLS) etc.) [44,114].

During this Ph.D. study, software by the instrument vendor (Agilent Technologies) was used for data processing e.g. Mass Hunter Qualitative and Quantitative analysis. For targeted analysis, qualitative analysis software was mainly used for reviewing chromatograms, peak finding, evaluation of baseline etc.. For the multitargeted approach this software was used for putative identification of the known unknowns extracted from different matrices. In this respect data were processed using the find by formula algorithm where the empirical formula was used to find matching masses in the data. Quantitative analysis program was used for creating calibration curves and quantification. For the multitargeted approach this program was used fast screening of compounds between different samples.

Although statistical data analysis was not performed during this Ph.D., it is worth mentioning the application of the commonly used statistical tools for mass spectrometry data analysis.

Clustering (shown as a heat map) is a statistical method that involves dividing observed datasets into several subclasses or clusters. This method was use by *Hou et al.* [34] to investigate the impact of decreased NADH levels to the concentration of the metabolites

involved in the central carbon metabolism when grown on glucose and ethanol. The statistical analysis showed that the growth conditions had a bigger impact on the metabolic profiles than the perturbations. Furthermore when grown on glucose the concentration of many glycolytic and TCA metabolites was increased compared to when grown on ethanol.

Principle component analysis (PCA) is another statistical tool used to discover and visualize important differences between groups of samples. PCA together with clustering were used to investigate the effects of i) loss of cytosolic superoxide dismutase function and ii) chemical-induced oxidative stress as well as iii) the metabolic profiles of different fly species of *Drosophila melanogaster* [56]. Clustered heat map was also used by *Brauer et al*. [115] to investigate the changes in the concentrations of the intracellular metabolites in *S. cerevisiae* and *E. coli* after glucose and nitrogen starvation.

2. Results and discussion

2.1 Targeted metabolomics

2.1.1 Optimization of liquid chromatography QqQ method

The primary focus in developing the presented method was qualitative and quantitative analysis of various intracellular metabolites. Determination and optimization of the compound specific multiple reaction monitoring (MRM) transitions (**Paper 1**) was the first step in the establishment of the method. Most of the intracellular metabolites are phosphorylated or carboxylated, producing mainly $[M-H]^-$ ions. Therefore the observed similarities in their fragmentation patterns were expected. Phosphorylated compounds generated $[H_2PO_4]^-$ and $[PO_3]^-$ ions as the most intense fragments, while loss of CO_2 was observed in the spectra of the carboxylated compounds. Thus, to avoid false positive results due to non-specific transitions, specific fragments (e.g. corresponding to purine or pyrimidine groups of the nucleotides) were used when possible. This was especially necessary when faced with analysis of compounds with the same elemental composition such as ATP and deoxyguanosine 5'-triphosphate (dGTP).

Taking into consideration the anionic nature of the intracellular metabolites several possibilities were tested for their chromatographic separation. Mixed-mode chromatography (Acclaim® Trinity™ P1) combines multiple retention mechanisms such as cation exchange, anion exchange and reverse phase mechanism of retention. The possibility for separation of compounds with different functionalities made this type of chromatography very attractive and it was therefore tested for separation of the intracellular metabolites with sugar phosphates as model compounds. Retention on the column, was obtained (Figure 9), but both a salt and a pH gradient did not give a sufficient separation of these compounds.

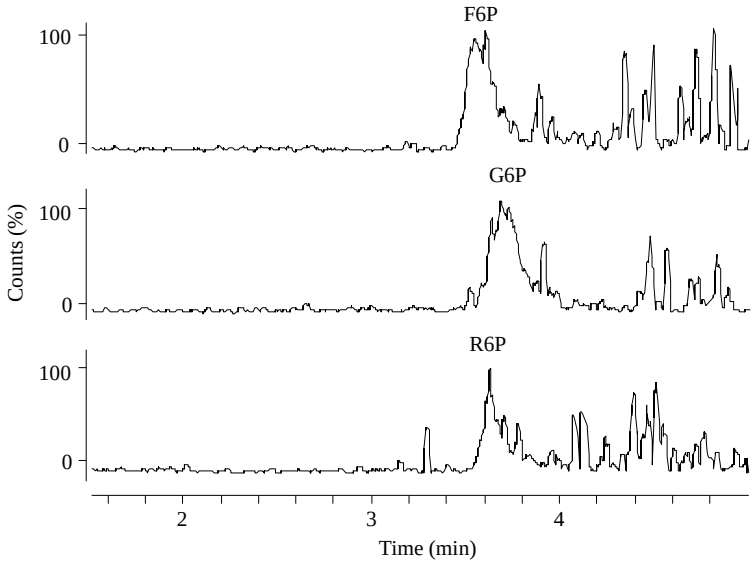


Figure 9. Chromatograms showing similarities in RT between the sugar phosphates F6P, G6P and R5P obtained using mixed mode LC-QqQ. Injection volume: 1µl; Concentration: 1µg/ml. Column: Acclaim Trinity P1. Eluent A: 50 mM ammonium formate pH 3.5; Eluent B: ACN. Gradient: 0-3 min 70-0 % B; 3-4 min 0 % B; 4-4.1 min 0-70 % B; 4.1-5 min 70 % B. F6P, fructose 6-phosphate; G6P, glucose 6-phosphate; R5P, ribose 5-phosphate.

Furthermore the high concentration of salt (50 mM ammonium formate) necessary to elute the sugar phosphates was not an ESI-MS friendly mobile phase (caused contamination of the ion-source resulting in suppression of the MS signal). This was an additional reason not to proceed with further optimizations of the mixed mode based chromatographic separation.

Interfacing IC with a Bruker Q-TOF to investigate the applicability of this type of liquid chromatography for separation of the intracellular metabolites was a part of a master project in which I was involved as a supervisor [116]. Compounds ranging from nucleotides, sugar phosphates, organic acids and coenzymes were taken as model compounds. The main issues encountered when using IC for separation were related to the instability of some of the intracellular metabolites under the extreme high pH used in the eluent. Multiple unknown chromatographic peaks were detected in the chromatograms of the redox compounds. Flavin adenine dinucleotide (FAD) showed to be degraded to AMP under the high pH conditions,

while hydrolysis of acetyl coenzyme A (Ac-CoA) resulted in only one peak in the chromatogram that corresponded to CoA. Dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (G3P) showed to be unstable during the chromatography as well. Phosphate was the main peak detected in the DHAP standard, while two peaks were detected in the chromatogram of G3P, one corresponding to G3P and the other one to phosphate.

Due to the lack of time no further experiments were carried out using the IC-Q-TOF MS technique. Despite the issues regarding the instability of some of the intracellular metabolites, IC showed a promising potential as a technique for analysis of intracellular metabolites. It gave good retention and separation of the various phosphorylated and carboxylated metabolites and showed to be suitable for analysis carboxydrates as well. In future it can be considered as a complementary technique to the IP-RP.

2.1.1.1 Ion-pair chromatography

Ion-pair chromatography has previously been proven to be a good separation technique for phosphorylated and carboxylated compounds [17,57,62,63,80] and was also used during this study. **Paper 1** outlines the main part of the work related to the development of the ion-pair chromatographic method. Several parameters such as type of ion-pair reagent, concentration of ion-pair reagent and acetic acid, pH, organic solvent and column chemistry were previously shown to influence the retention of the anionic compounds when ion-pair chromatography was used [17]. Therefore during the development process these parameters had to be tested.

Volatile alkylamines with diverse alkyl chain length such as triethylamine (TEA), DBA and TBA, were firstly tested in order to find the best compromise between the retention and separation of the metabolites of interest. Ten millimolar TEA gave slightly increased retention but not enough to be able to separate for example the sugar phosphates. Increasing the concentration of the TEA in the aqueous mobile phase (from 10 to 35 mM) and the

addition of the ion pair reagent into the organic mobile phase did not increase the RT of for example G6P. Neither the incorporation of an isocratic part at the beginning of the run nor the lowering of the percentage of organic solvent during the run improved the retention. The length of the alkyl chain was considered to be responsible for the poor retention of the metabolites, therefore two other ion-pair reagents with longer hydrophobic chains were tested: DBA and TBA. Bothe DBA and TBA resulted in improved retention. In the case of DBA the intensity of the nucleotides was decreased when compared to TBA. In addition better separation for AMP and deoxyguanosine 5'-monophosphate (dGMP) was obtained with TBA than with DBA when using the same gradient. Therefore TBA was chosen for further optimizations. The changes of the TBA concentration while keeping the ratio between the ion-pair reagent and the acetic acid unchanged, in order to keep the pH constant, gave interesting results on the retention of the different groups of compounds tested (Figure 10).

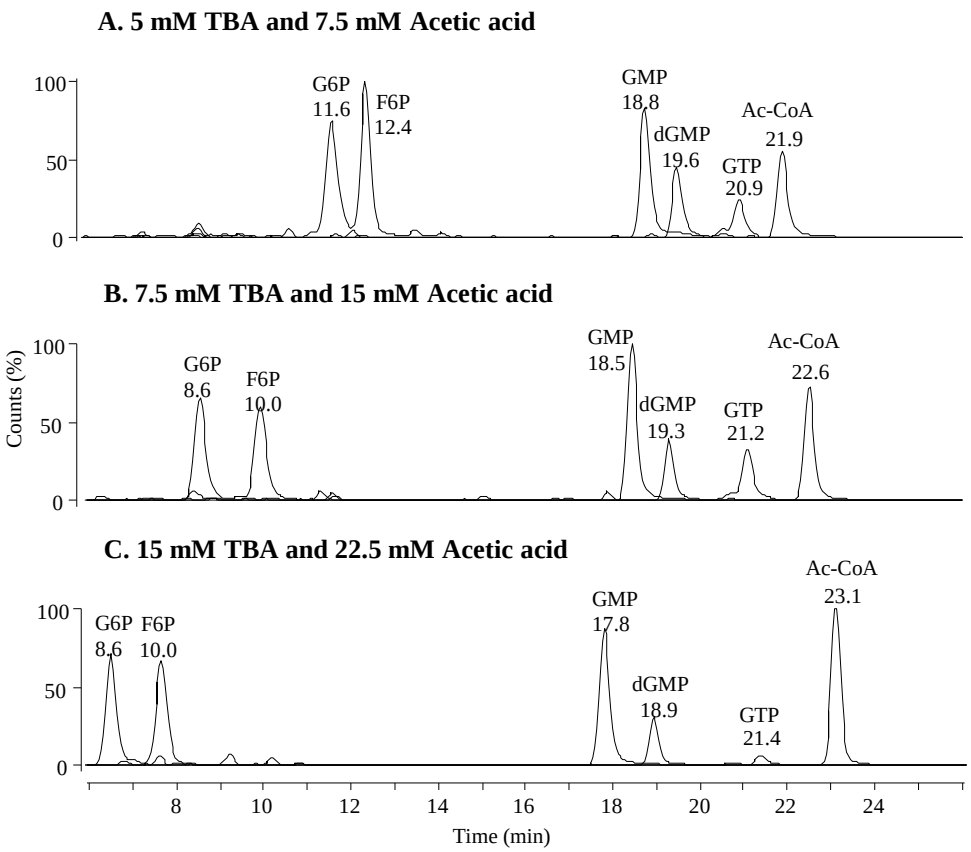


Figure 10. Chromatograms showing the effect of increased concentration of TBA in the eluent on the retention of different intracellular metabolites, while keeping the ratio between TBA and acetic acid constant. Column: Luna 2.5 µl C18 (2)-HST (100 x 2 mm). Eluent B was 90 % MeOH containing the same concentration of TBA and acetic acid as eluent A, respectively. Gradient: 0-5 min 0 % B; 5-10 min 0-2 % B; 10-11 min 2-9 % B; 11-16 min 9 % B; 16-18 min 9-25 % B; 18-19 min 25-50 % B; 19-23 min 50 % B; 23-24 min 50-0 % B; 24-29 min 0 % B. GMP, guanosine 5'-monophosphate; GTP, guanosine 5'-triphosphate; Ac-CoA, acetyl coenzyme A.

In the case of the early eluting compounds such as F6P and G6P, decrease of the retention was observed with the increase of the TBA concentration in the eluent. Guanosine 5'-monophosphate (GMP) and dGMP showed only slight decrease in the retention by increasing the concentration of the ion-pair reagent, while guanosine 5'-triphosphate (GTP) showed slight increase. The late eluting compound acetyl-coenzyme A (Ac-CoA) showed increase in the retention by increasing the TBA concentration.

As described in **Paper 1** the concentration of the acetic acid, consequently the pH, had also an influence on the retention and separation of the intracellular metabolites. Due to the competitive effect of the acetate ion, by increasing the concentration of the acetic acid, decrease of the retention was observed. Furthermore, by changing the pH from 5.5 to 7.5 changes in the intensity ratio between the monophosphates, diphosphates and triphosphates occurred. At lower pH the MS signal intensity of the monophosphates and diphosphates was higher than the triphosphates while an opposite trend was observed with higher pH. However taking into account the investigations made with changing the concentration of TBA and acetic acid, 10 mM TBA and 10 mM acetic acid was found to give a reasonable compromise between elution time, resolution and sensitivity.

Isopropanol and MeOH were tested as organic solvents. Due to the stronger elution power of isopropanol, the separation of the compounds was affected and therefore MeOH was chosen for further optimizations. In addition, different reverse phase columns were also tested such as Phenomenex Luna C_{18}(2)-HST, Onyx Monilithic C_{18}, Agilent Zorbax extended C_{18} and Poroshell 120 Phenyl-Hexyl. The effect of the column chemistry was mostly on the peak shape and the spreading of the compounds with Poroshell 120 Phenyl-Hexyl column giving the best performance.

It should be mentioned that, a newly installed column on the system was equilibrated overnight with the eluent containing the TBA. Furthermore, it was noted that when using the 36 min. gradient (**Paper 1**) shifting of retention times occurred only when changing to new eluents. There was a shift in the RT for the compound eluting in the middle of the gradient e.g. monophosphorylated compounds. This was assumed to be due to the very slow gradient in the region where the monophosphorylated compounds were usually eluting.

Furthermore, the effect of the TBA addition in the injection vial on the analysis of the nucleotides was tested as well. By injecting a mixture of the nucleotides with and without

TBA, at least 2x increase in the signal intensity was observed when injecting the samples with TBA (Figure 11).

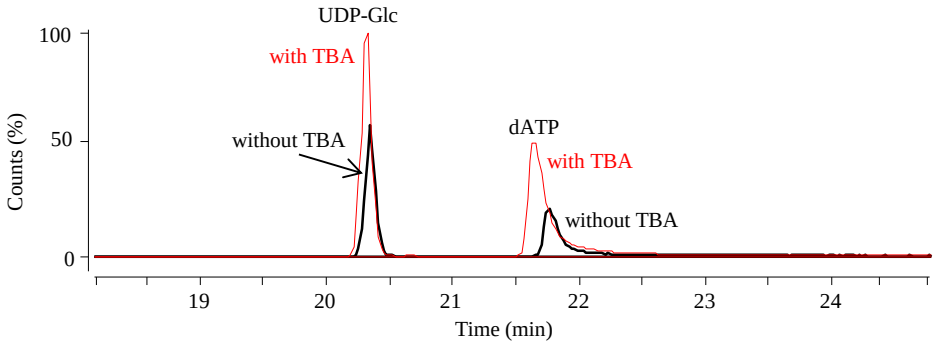


Figure 11. Chromatograms showing the effect of addition of TBA in the injection vial. Injection volume: 10 µl; Concetration: 10 µg/ml. Column: LUNA 2.5µm C18(2)-HST. Eluent A:10 mM TBA and 10 mM acetic acid, eluent B: 90 % MeOH containing 10 mM TBA and 10 mM acetic acid. Gradient: 0-5 min 0 % B, 5-10 min 0-2 % B, 10-11 min 2-9 % B, 11-16 min 9 % B, 16-18min 9-25 % B, 18-19 min 25-50 % B, 19-23 min 50 % B, 23-24 min 50-0 % B, 24-34 min 0 % B. UDP-Glc, uridine 5'-diphosphate; dATP, deoxyadenosine 5'-triphosphate.

This has been explained by a displacement mechanism, where the common potassium and sodium adducts are displaced by alkylamine adducts which increases the signal of the singly charged molecular ions $[M-H]^-$ [17, 117].

In the case of the NADH and NADPH, it was noticed that when injecting pure standards of these two compounds, peak of NAD^+ was detected in the NADH standard and $NADP^+$ was detected in the NADPH standard. This indicated that oxidation was happening in the vial and was noted when injecting the standards both with and without TBA in the injection vial. As previously reported in the literature [87] ammonium acetate pH 8 has been shown to improve the stability of NADPH and reduce the oxidation. In order to investigate this, the standards of NADH and NADPH were prepared in 5 mM ammonium acetate pH 8, which indeed reduced the oxidation of these two compounds. However the addition of TBA in the injection vial was proven to increase the sensitivity of the nucleotides. Therefore 50 µg/ml solutions of NADPH and $NADP^+$ was prepared by diluting the ammonium acetate stock solutions (pH 8) of these

two compounds into eluent A which contained TBA. It was noted that the addition of TBA did increase the sensitivity especially for NADPH, while the intensity of NADH was practically not affected. Furthermore, there was a minor increase of the NAD^+ LC-MS peak in the NADH standard, while no peak of $NADP^+$ was detected in the NADPH standard. However further investigations are necessary in order to check the long term stability of the NADH and NADPH solutions prepared in the ammonium acetate as well as the dilutions of these stocks into TBA.

Within the intracellular metabolites, there are many compounds (sugar phosphates, ATP/dGTP, sugar nucleotides etc.) with the same elemental composition which has been a challenge during the method development. These compounds can often be indistinguishable by MS due to their identical elemental composition as well as similar fragmentation pattern. Therefore the gradient used during the analysis was crucial for the separation especially for the sugar phosphates. The 5 min isocratic run at the beginning of the 36 min gradient given in **Paper 1** was necessary for achieving the separation of the sugar phosphates. By introducing even 5 % of the organic phase at the beginning of the gradient, the separation of the sugar phosphates was impaired. In addition, increasing the percentage of the organic phase up to 100 % and keeping it for 1.5 min was necessary for a complete elution of the more retained compounds.

In the case of the nucleotides, some of the isomers such as AMP/dGMP were chromatographically separated, while that was not the case for ATP/dGTP. Due to the fact that ATP and dGTP were chromatographically not separated, specific fragments that correspond to the guanine or adenine moiety were chosen in order to be able to quantify these compounds.

The thirty six minutes gradient allowed a good separation of many of the isomers, but in general it was considered as time consuming and not very practical when long lists of

samples had to be analyzed. Therefore, compromise between the analysis time and the separation of the nucleotides was found by a 19.5 min gradient (**Paper 5**). However this gradient could not be used for separation of the sugar phosphates. The separation of the isomers was also considered as a parameter for the performance of the column. The impaired separation for example of the sugar phosphates was a sign that the column had to be changed.

The retention of the compounds when using ion-pair chromatography depends on the number of charged groups present in the molecule that could interact with the ion-pair reagent. In the case of the phosphorylated and/or carboxylated compounds, correlation between the retention time (RT) and the number of the phosphate and/or carboxylic groups was observed. The order of elution for example of the phosphorylated nucleotides was monophosphates < diphosphates < triphosphates. The same was observed for the carboxylated compounds. Understanding the mechanism of the chromatographic retention allowed prediction of the RT of the known unknowns and helped in their identification in the sample extract although standards for these compounds were not available in house (**Paper 6**).

The fact that the reversed phase mechanism of action is also involved in the retention was shown by the analysis of the amino acids (**Paper 6**). The aromatic amino acids, tyrosine (RT 1.3 min) and phenylalanine (RT 2.2 min) were retained using ion-pair chromatography (Figure 12), although their overall charge was 0.

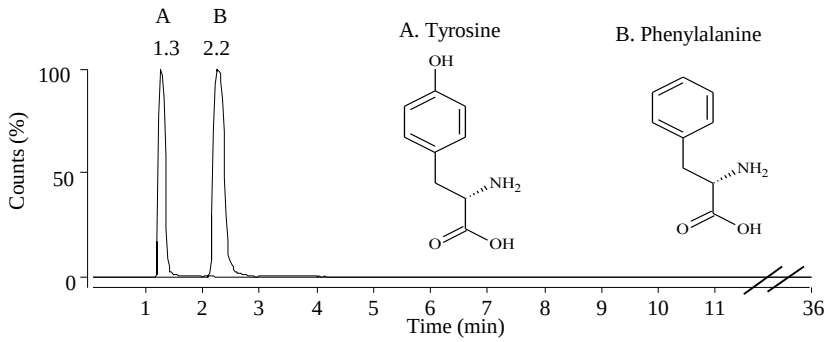


Figure 12. Chromatogram showing the reversed phase mechanism of retention during the IP-RP by retaining the tyrosine (RT 1.3 min) and phenylalanine (RT 2.2 min) which overall charge was 0. (Dead volume = 0.8 min). Column used: Poroshell 120 Phenyl-Hexyl. Eluent A:10 mM TBA and 10 mM acetic acid, eluent B: 90 % MeOH containing 10 mM TBA and 10 mM acetic acid. Gradient: 0-5 min 0 % B, 5-10 min 0-2 % B, 10-11 min 2-9 % B, 11-16 min 9 % B, 16-24 min 9-50 % B, 24-28 min 50 % B, 28-28.5 min 100 % B, 28.5-30 min 100 % B, 30-30.5 min 100-0 % B, 30.5-36 min 0 % B.

This could be explained by the reversed phase interactions between the aromatic part of the amino acids and the phenyl hexyl groups from the stationary phase. Although the column was covered by the ion-pair reagent, the reverse phase mechanism of interaction was also responsible for the retention. Furthermore the tyrosine was eluting earlier due to the presence of the polar hydroxyl group.

Summary of the tested conditions during the optimization of the ion-pair separation of the intracellular metabolites and the results obtained is given in Table 5. Furthermore, a list of standard metabolites for which detection was optimized is given in Supplementary material.

Table 5. Summary of the tested conditions and the results obtained from the optimization of the ion-pair chromatography.

Tested conditions	**Result**
Type of ion pair reagent	RT increases by an increase of the number and the length of the alkyl chains (TEA<DBA<TBA) • TEA - not enough RT to achieve good separation • DBA- lower intensity of the nucleotides when compared to TBA • TBA - gives better separation of the isomers AMP and dGMP than DBA
Increasing the concentration of TBA while keeping the pH constant	• Early eluting compounds - RT decreases • Compounds eluting in the middle of the gradient - RT slightly affected • Late eluting compounds - RT increases
Changing pH while keeping the concentration of TBA constant	• pH 5.4: NMPs and NDPs higher intensity than NTPs[1] • pH 7.7: NMPs and NDPs lower intensity than NTPs
Column chemistry	Affects peak shape and separation
Organic solvent	• Isopropanol - stronger elution power than MeOH • Better compromise between the retention and separation with MeOH than isopropanol
Addition of TBA in the injection solvent	In general increases the intensity of the analytes
Isocratic run at the beginning of the run	Necessary to separate the sugar phosphates
Keeping 100 % of organic phase during the gradient	Necessary for complete elution of the more retained compounds
Column equilibration	Overnight equilibration with the eluent containing the ion-pair reagent is necessary

[1]NMPs - nucleotide monophosphates; NDPs – nucleotide diphosphates; NTPs – nucleotide triphosphates

2.1.2 Establishment of quenching and extraction procedures

This section outlines the work done during this Ph.D. study related to the testing of different sample preparation methods for yeast, bacteria, filamentous fungi and mammalian cells. ECR of 0.80-0.95 was taken as criteria for a successful quenching. The organisms for which several different quenching and extraction procedures were tested during this Ph.D. study as well as the aim of the studies are given in Table 6.

Table 6. Quenching and extraction procedures tested, the obtained energy charge ratio and the aim of the study of the different organisms during this Ph.D.

Organism	Aim of the study	Quenching	Extraction	ECR (SD)[1]	Metabolites measured
Saccharomyces cerevisiae	Is the difference in the production of vanillin-β-glucoside between two yeast strains related to the differences in their intracellular metabolite pools?	1. -40 °C MeOH/ H_2O (60/40, v/v)	1.1. Boiling EtOH/H_2O (75/25, v/v) 1.2. MeOH/CH_3Cl (2:1, v/v)[2]	1. 0.91 (0.01) 1.2. 0.91 (0.006)	ATP, ADP, AMP, redox cofactors, UDP, UDP-Glc and UTP
Lactococcus lactis	Dispersive solid phase extraction of nucleotides	10 M HCOOH	3x freeze/thaw	0.97(0.0001)	Nucleotides
Ruegeria mobilis	Are shifts between motile and sessile life correlated to intracellular concentrations of c-diGMP?	1. 10 M HCOOH 2. cooling on ice	1. 3x freeze/thaw 2. Boiling EtOH/H_2O (75/25, v/v)	- -	c-diGMP
Microbispora corallina[a]	Is the improved production of lantabiotic correlated to the changes in the energy metabolism?	1. 10 M HCOOH 2. -40 °C MeOH/glycerol (60/40, v/v) 3. -40 °C MeOH/glycerol (40/60, v/v) 4. Phosphate buffered saline (PBS) 5. -40 °C MeOH/10 mM HEPES (60/40, v/v) 6. -40 °C MeOH/ H_2O (60/40, v/v)	1. 3x freeze/thaw 2. Boiling EtOH/H_2O (75/25, v/v) 3. Boiling EtOH/H_2O (75/25, v/v) 4. CH_3COONH_2 and bead beater 5. Boiling EtOH/H_2O (75/25, v/v) 6.1. Boiling EtOH/H_2O (75/25, v/v) 6.2. MeOH/CH_3Cl (2:1, v/v)	1. 0.80 (0.004); 0.28 (0.1) 2. 0.06 (0.01) 3. 0.06 (0.02) 4. 0.16 (0.03) 5. 0.62 (0.03) 6.1. 0.58 (0.07) 6.2. 0.87 (0.01)	ATP, ADP, AMP, redox cofactors
Streptomyces coelicolor	The effects of altered redox and energy levels on the production of two antibiotics.	1. 10 M HCOOH 2. -40 °C MeOH/ H_2O (60/40, v/v)	1. 3x freeze/thaw 2.1. Boiling EtOH/H_2O (75/25, v/v) 2.2. MeOH/CH_3Cl (2:1, v/v)	1. 0.54 (0.01) 2.1. 0.52 (0.12) 2.2. 0.89 (0.03)	ATP, ADP, AMP, redox cofactors
Chinese hamster ovary cells (CHO)	Are the differences in erythropoietin productivity between two clones related to the difference in their central carbon metabolism?	Ice cold 0.9 % (w/v) NaCl	MeOH and ACN	0.92 (0.02)	Glycolytic intermediates, ATP, ADP, AMP, redox cofactors

[1] The numbers in the brackets indicate standard deviation (SD)
[2] This extraction method was applied on yeast but not during the vanillin-β-glucoside project

[a] The *Microbispora corallina* project was called LAPTOP and was funded from the European Commission contract no. 245066 for FP7-KBBE-2009-3.

Table 6. Continued

Organism	Aim of the study	Quenching	Extraction	ECR (SD)	Metabolites measured
Aspergilus nidulans	Testing different quenching and extraction procedures (master thesis project)	1. -40 °C MeOH/ H_2O (40/60, v/v)	1.1 Centrifugation; MeOH/H_2O (40/60, v/v), 3x freeze/thaw	0.79 (0.01)	ATP, ADP, AMP
		2. 0°C 0.9% (w/v) NaCl	1.2 Filtration, MeOH/H_2O (40/60, v/v), 3x freeze/thaw	0.24	
			1.3 Centrifugation, MeOH/CH_3Cl (2:1,v/v)	0.80 (0.04)	
			1.4. Filtration, MeOH/CH_3Cl (2:1,v/v)	0.85(0.02)	
			2.1 Centrifugation; MeOH/H_2O (40/60, v/v), 3x freeze/thaw	0.79 (0.08)	
			2.2 Filtration, MeOH/H_2O (40/60, v/v), 3x freeze/thaw	0.24	
			2.3 Centrifugation, MeOH/CH_3Cl (2:1,v/v)	0.85 (0.02)	
			2.4. Filtration, MeOH/CH_3Cl (2:1,v/v)	0.84	

Yeast. Establishment of quenching and extraction procedures for *S. cerevisiae* has been a part of the work during this thesis. *S. cerevisiae* is a widely studied organism [27,34-37,79-81]. One of the most commonly applied quenching method for this organism is -40 °C MeOH/ H_2O (60/40, v/v). In order to check the leakage during the cold MeOH quenching, four different types of *S. cerevisiae* samples were quenched and extracted: whole broth, quenching supernatant, culture filtrate and the biomass. The results are shown in Figure 13.

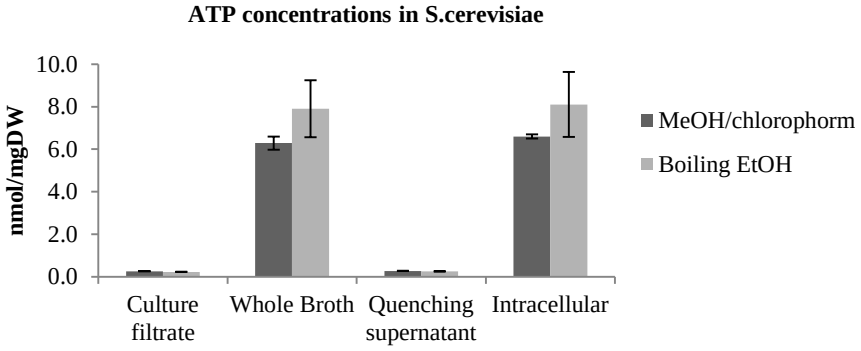


Figure 13. Concentrations of ATP measured in the culture filtrate, the whole broth, quenching supernatant and in the biomass of *S. cerevisiae* using -40 °C MeOH/ H_2O (60/40, v/v) as a quenching combined with boiling EtOH/H_2O (75/25, v/v) or MeOH/chloroform (2:1, v/v) as extraction.

The low concentration measured in the quenching supernatant indicated that the leakage percentage was very low. This was confirmed with the similar concentrations measured intracellular and in the whole broth.

This method was applied in two different studies (**Paper 4 and 5**) in combination with two different extraction methods: boiling ethanol and MeOH/chloroform. In both cases high ECR was obtained (in the range of 0.80-0.95).

Bacteria. As previously mentioned in the introduction part, bacteria are more prone to leakage during the quenching when compared to for example yeast [82,91]. Thus approaches where the cells are not divided from the medium were tested firstly. Change of the pH by addition of 10 M formic acid followed by 3x freeze thaw was shown to work very well for

quenching of *L. lactis* (**Paper 1**), giving an ECR of 0.970±0.001. Since the cells were not separated from the growth media, purification of the samples using dispersive solid phase extraction (DSPE) with charcoal was performed. This was done in order to reduce the matrix effects from the growth media and to separate the nucleotides from the other non-aromatic intracellular compounds. The retention of the nucleotides on the charcoal was based on the interaction of the aromatic ring electrons from the purine or pyrimidine moiety of the nucleotides with the π electrons form the charcoal. *Jendresen et. al.* [92] used high pH in combination with ethanol to elute the nucleotides from the charcoal. During this Ph.D. study other organic modifiers (MeOH, isopropanol and acetonitrile) as well as ion-pair reagent (DBA) were tested in order to investigate the possibilities of improvement of the recovery of the nucleotides. It was found that acetonitrile in combination with high pH gives the best elution of the nucleotides from the charcoal. The quenching using low pH proved to be efficient for quenching of *L. lactis* and was therefore tested for the other bacteria of interest during this Ph.D. study such as *Ruegeria mobilis* (*R. mobilis*), *M. corallina* and *S. coelicolor*. When the formic acid quenching method was applied on *R. mobilis* followed by 3x freeze-thaw cycles and charcoal sample purification, no c-diGMP was detected in the extract. The reason for this was considered to be the low intracellular concentrations of c-diGMP that might be lost during the sample preparation. Furthermore, strong matrix suppression was observed even after the purification of the samples using DSPE with charcoal due to the complex media used for growth. This was assumed to be due to the unspecificity of the charcoal as a sorbent. Many of the media components were not removed after the purification and were concentrated with the evaporation. After the injection of the samples into the LC-MS, deposition of the salts on the ion-source was observed that caused clogging of the ion source. Therefore an alternative sample preparation method was tested by cooling the cells on

ice followed by addition of boiling EtOH/H_2O (75/25, v/v) which resulted in more clean samples and detection of the intracellular c-diGMP in *R. mobilis* (**Paper 2**).

As previously mentioned, formic acid was also tested for quenching of the metabolic activities in *M. corallina* and *S. coelicolor* but without acceptable results. For *M. corallina* the main problem was the reproducibility of the acceptable values for the ECR (Table 6). Different values for the ECR were obtained on different days and were in the range between 0.28-0.80 showing that the method was not reproducible. The reason for the non-reproducible results was speculated to be morphology related. *M. corallina* is filamentous bacteria that can exist in more dispersed or pelleted form depending on the nitrogen source used for growth. When sampling from the fermentors with pelleted growth (Figure 14), the cells settled on the bottom of the spin tube immediately after the sampling.

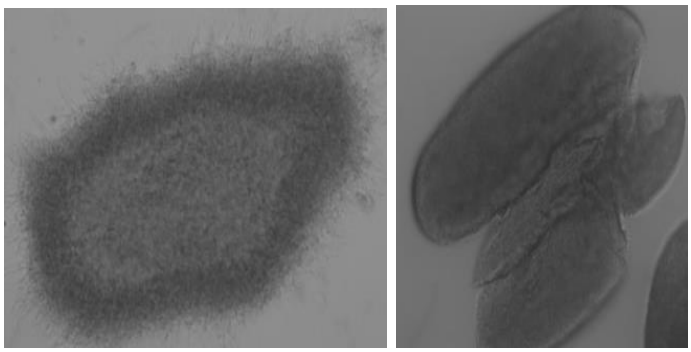


Figure 14. Microscopic pictures of *M. coralline* pellets[2] (magnified 60x).

This together with the pelleting itself was speculated to result in non-reproducible dispersion of the formic acid around the cells thus non reproducible quenching (ECR between 0.28-0.60). When existing in more dispersed phase better ECR was obtained but it was not reproducible (ECR in the range 0.64-0.80). It was speculated that this was due to grouping of the filaments.

S. coelicolor is also a filamentous bacteria thus clump formation by the filaments was expected. Similar as for *M. corallina* the clump formation resulted in less effective quenching

[2] Pictures taken from Subir Nandy, PostDoc at the section for Eukaryotic Biotechnology, Department of Systems Biology, Technical University of Denmark and collaborator in the *Microbispora corallina* project.

thus low ECR quenching with formic acid as a quenching solution. The low pH applied resulted in ECR of 0.54 which was not within the acceptance criteria (ECR of 0.80-0.95). Therefore other quenching and extraction methods were tested.

As previously mentioned in the introduction part, glycerol combined with saline or MeOH has been shown to be a promising quenching mixture for microbial cells [84,85] and was therefore chosen for testing. However during this study, several problems were encountered with the glycerol based quenching solution: i) difficulties to determining leakage ii) longer time for separation of the cells from the supernatant by centrifugation; and iii) residues of the glycerol in the extract. The high amounts of glycerol present in the quenching supernatant could not be removed due to the high boiling point of glycerol. Therefore it was impossible to check the leakage during the quenching. Furthermore, the high viscosity of the saline/glycerol solution slowed down the decanting process and removal of the quenching supernatant from the cells. This was considered as a major drawback of this procedure due to the longer exposure of the cells to the quenching solution and therefore increasing the probability of leakage. Even exchanging the water with MeOH to decrease the viscosity and to allow operation at even lower temperatures did not improve the whole quenching process. Furthermore, the residues of the glycerol in the samples made them more viscose after the concentration causing problems in the analysis step (it was impossible to inject the samples into the LC-MS). Therefore the cells were washed twice with cold 0.9 % (w/v) NaCl in order to remove the glycerol residues. However, the ECR obtained was extremely low showing that the quenching was not successful. Due to the technical problems encountered, the MeOH/glycerol quenching was not further investigated as a quenching solution for *S. coelicolor*.

Since one of the aims of the studies for *M. corallina* and *S. coelicolor* was also measurement of the redox pairs, quenching and extraction methods that will prevent the

oxidation/reduction of the redox compounds were necessary. Therefore a sample preparation procedure previously reported by *Sporty et al*. [88] for accurate measurement of NAD^+ and NADH redox state was tested for *M. corallina*. Phosphate buffered saline was used for quenching the cells while the nitrogen saturated ammonium acetate used during the extraction was expected to prevent the oxidation of NADH to NAD^+. However the concentrations of NAD^+ and NADH obtained with this method were similar to the ones obtained with the other quenching methods. Furthermore, slightly higher concentration of $NADP^+$ and quite low ECR values were obtained with this method when compared to the other quenching and extraction procedures tested for *M. corallina*.

Kassama et al. [97] applied the commonly used quenching method for yeast (-40 °C MeOH/H_2O (60/40, v/v)) to quench the metabolism of another species from the *Streptomyces* genus but the authors did not present data regarding the leakage. Due to the difficulties to find a method that will give an acceptable ECR value for *M. corallina* and *S. coelicolor* the yeast quenching method was further tested although leakage was expected to occur. This was done in order to inspect if the low ECR obtained using the previous quenching methods, were due to physiology reasons or due to the method used for quenching. When testing the cold -40 °C MeOH/H_2O (60/40, v/v) quenching in combination with boiling EtOH as an extraction method ECR of 0.6 was obtained. This was unexpected result since it is well known that this type of quenching is able to stop the metabolic activities within a second and the problems are mainly related to the leakage. It was suspected that the high temperatures used during the extraction might cause conversion of ATP to ADP and AMP, leading to a low ECR. An additional attempt for improving the ECR value was done by changing the boiling EtOH extraction with MeOH/chloroform instead, which resulted in ECR values of around 0.87. In order to check the leakage the quenching supernatants were collected. Since the biomass pellet obtained after centrifugation was easily re-suspended in the quenching

solution, some of the biomass was lost during decanting. Thus there was a limited number of quenching supernatant samples that were free of the cells that could be used to check the leakage. However, the peak area of ATP, ADP and AMP detected in the supernatant was the same as in the biomass, indicating a severe leakage.

Since sever leakage was observed for *M. corallina* an attempt was made to decrease the possible leakage that might occur in the case for *S. coelicolor*. As previously reported by *Wellerdiek et al.* [118] the metabolic reactions in quenched *Corynebacterium glutamicum* were stopped at -20 °C. Furthermore, *Jonge et al*. [95] showed that by increasing the temperature of the cold MeOH from -40 to -25 °C, successful quenching of *Penicillium chrysogenum* was achieved with decreased leakage. Therefore the MeOH/H_2O (60/40, v/v) quenching method was tested using two temperatures: -40 and -25 °C both with boiling EtOH and MeOH/chloroform as an extraction methods. As for *M. corallina*, an acceptable ECR was observed only when MeOH/H_2O (60/40, v/v) was combined with MeOH/chloroform. Similar concentrations were obtained for ATP, ADP, AMP and the redox compounds despite of the quenching temperature used for extraction. In all cases ECR of 0.80-0.95 was obtained. To investigate if any significant leakage occurs during the quenching, the adenylates ATP, ADP and AMP were measured in the quenching supernatant. Leakage was observed despite of the temperature tested and there was not difference in the leakage percentage between the two different temperatures. The highest leakage was observed for AMP (Figure 15). The differences in leakage between the compounds was explain by the fact that relatively smaller molecules permeate the cell membrane more easily than the larger polar molecules [28,81,82].

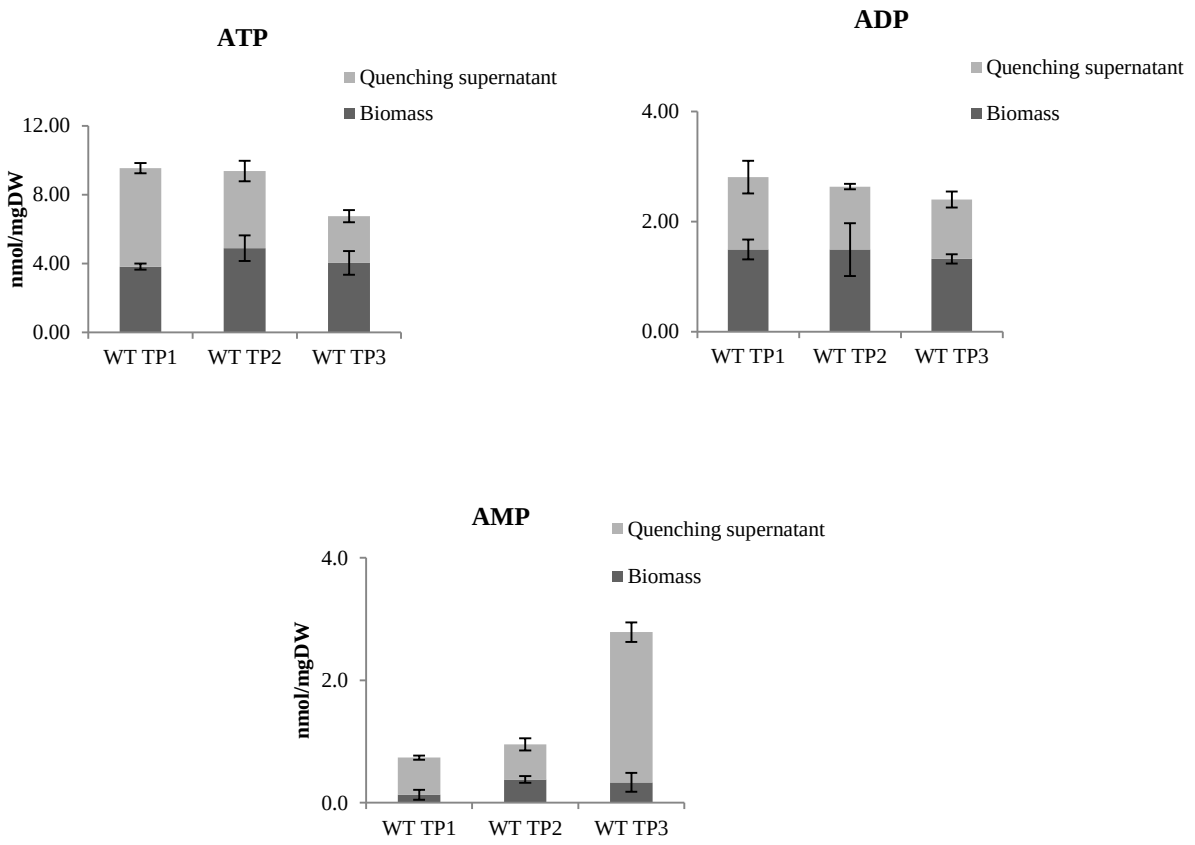


Figure 15. Concentrations of ATP, ADP and AMP measured intracellular and in the quenching supernatant from the WT strain of *S. coelicolor* at three time points during the exponential phase. The data presented are from the quenching at -40 °C. The error bars indicate standard deviations from three technical replicates.

The high concentration of AMP measured in WT time point (TP) 3 (Figure 15) was probably due to quenching problems. This was also reflected in the lower concentration of ATP measured in the same sample.

Mammalian cells. The work related to the mammalian CHO cells was a part of a master project which main goal was to investigate if the additional metabolic burden induced by recombinant protein production affects the central carbon metabolism (**Paper 3**) (the analysis of the intracellular metabolites in this project was supervised by me). *Dietmar et al.* [24] showed that ice cold 0.9 % (w/v) NaCl did not to damage the membrane of the mammalian cells and was capable of arresting the cell metabolism. Therefore ice cold 0.9 % (w/v) NaCl as a quenching solution was combined with cold MeOH and ACN/H_2O (50/50, v/v) as

extraction solvents which resulted in ECR values ranging from 0.8 to 0.9. No further optimization of the sample preparation was done due to the limited time that was available for this project.

Filamentous fungi. Testing different quenching and extraction methods for filamentous fungi in this case, *Aspergillus nidulans*, was a part of a master project [119] in which I was involved as a supervisor. For this purpose, two types of quenching procedures: i) -40 °C MeOH/H_2O and ii) 0 °C 0.9% (w/v) NaCl were combined with two extraction procedures i) MeOH combined with freeze thaw in liquid nitrogen and ii) MeOH/chloroform as given in Table 6. In addition both centrifugation and filtration were tested for separating the cells from the quenching solution.

The reason for testing cold MeOH as a quenching method was due to the wide applicability of this method for quenching the metabolism of various microorganisms including *Penicillium chrysogenum* [95]. On the other hand, the isotonic water was considered to be a less aggressive quenching method consequently less leakage was expected during the quenching. Furthermore isotonic water has previously been shown to be a good quenching method for mammalian cells (**Paper 3**).

In general, the problems that occurred were related to the technique itself and the leakage. The incomplete separation of the biomass from the quenching supernatant, when centrifugation was used, made the measurements of the leakage unreliable. Therefore all the methods that included centrifugation as separation were not taken into account for further optimization. From the methods that used filtration as a separation technique, the best energy charge ratio (~ 0.8) was observed with both the -40 °C MeOH/H_2O (40/60, v/v) and isotonic water combined with MeOH/CH_3Cl (2:1, v/v) as an extraction method. The -40 °C MeOH/H_2O (40/60, v/v) was further investigated for leakage by comparing the amounts of the ATP, ADP and AMP detected in the quenching supernatant with those measured in the

biomass. Leakage was observed and was more pronounced for AMP and ADP than for ATP. The amounts of ADP and AMP found in the supernatant were around 30 % of the sum of the amounts detected in the supernatant and the biomass while for ATP it was 3 %. Furthermore, the concentrations of AMP and ADP in the supernatant were 10 times higher than those detected for ATP. This was considered to be due to the higher diffusion rate of smaller metabolites with lower net charge.

This study showed that isotonic water can also be used for quenching of filamentous fungi. However, further investigations regarding the reproducibility of this method and the leakage percentage will be beneficial.

The work covered in this section once more demonstrates the challenges related to the establishment of quenching and extraction methods, especially for bacteria and filamentous fungi. Furthermore, the transfer of sample preparation methods from the literature to the lab is not an easy task due to: i) the difference in the microorganism of interest, ii) sampling technique, iii) quenching equipment and iv) operator. Validation of the quenching method is necessary for the different organisms in order to obtain reliable data that represent the metabolic state of interest. When the cells are not separated from the quenching supernatant, special care needs to be taken due to the contaminants introduced to the sample, for example from the growth media. As shown for *R. mobilis*, the analysis might be significantly impaired by the sample preparation method, thus minimal growth media are preferred over complex media when the quenching and extraction is combined. Yeast has shown to be the easiest organism to work with however leakage during the quenching does occur as previously reported by other groups [81].

2.1.3 Method validation and quantification

In this section, the approaches for preparing the calibration curves and the quality control samples, used for quantification and validation are discussed. All the validation and

quantification experiments during this Ph.D. were based on i) the standard addition or ii) the external calibration approach (either matrix matched or calibrants prepared in the neat solution). In each of the approaches commercially available SIL-IS were used.

Paper 1 discusses the challenges related to the quantitative measurement of the nucleotides in *L. lactis*. In this paper the validation and the quantitative measurements were performed using: i) the standard addition approach as suggested by *Tsikas et al*. [108] and ii) the external calibration approach where the growth media was used as a matrix for spiking and preparing the calibration curves and the quality control samples. Furthermore SIL-IS were used in both approaches which showed to be very important for improvement of the linearity, accuracy/recovery and precision of the analytical method.

The two major challenges when the standard addition approach was used were: i) the choice of the spiking concentrations and ii) the non-linearity of the calibration curves. *Tsikas et al.* [108] approach was used to determine the spiking concentration levels and the amounts added were approximately ranging from 50-250 % of the analyte concentration determined in the matrix, using an external calibration approach and corrected for the recovery.

A narrow linear range, due to the detector saturation at the high concentration end of the calibration curve was especially prominent for the nucleotides that were present in a high concentration in the spiking matrix. The addition of SIL-IS improved the linearity (Figure 16) and resulted in calibration curves with correlation coefficients higher than 0.99, for most of the nucleotides.

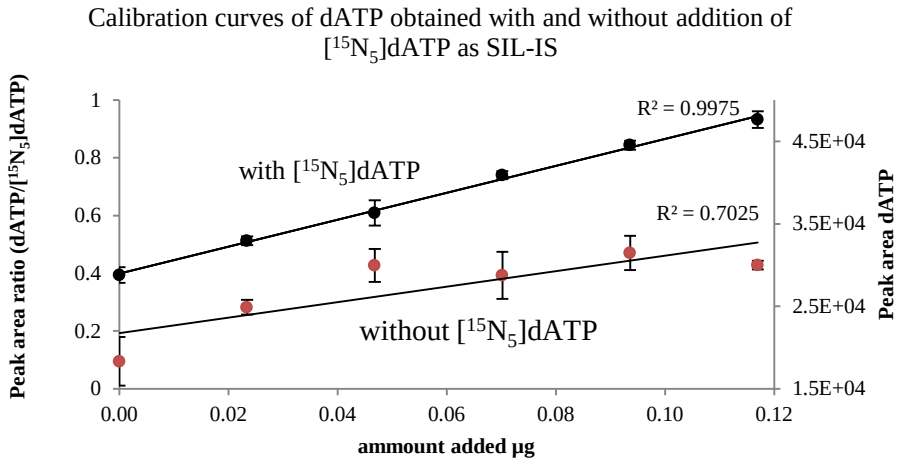


Figure 16. Graph showing the comparison of the calibration curves for dATP obtained using the standard addition methods with and without addition of the SIL-IS [$^{15}N_5$]dATP.

For the compounds such as ATP, GTP and UTP, that were present in high amounts in the matrix, it was noted that the intensity of their SIL-IS decreased by increasing the concentration of the coeluting non-labeled analyte. This was considered to be due to an ion-suppression of the SIL-IS signal by the highly concentrated coeluting non-labeled analyte. Consequently non-linear calibration curve was obtained even when using the SIL-IS. The problem of suppression of the SIL-IS by the coeluting non-labeled analyte has previously been described by *Sojo et al.* [120] and was solved by addition of higher amounts of the SIL-IS. However, for the compounds for which no SIL-IS was available, linearity improvement of the calibration curve was not possible. In the case of UDP-Glc the saturation of the detector due to the high amounts of this compound present in the spiking matrix resulted in accuracy and precision out of the acceptance criteria (acceptance criteria: accuracy within 80-120%; precision RSD ≤ 20 %). Appropriate dilutions of the spiking matrix could solve the problem with the non-linearity of UDP-Glc calibration curve. However this would require separate validation and quantification for this compound, which shows that the validation and quantification of the intracellular metabolites is less straightforward and might require additional steps for obtaining reliable results.

In the case of the external calibration approach (with the growth media as a spiking matrix), the major problem was considered to be related to the difference in the matrix effects between the matrix used to prepare the calibration standards and the matrix in which the compounds of interest needed to be quantified. In the standard addition approach the matrix effects were corrected. However, this was not the case for the external calibration approach, since the matrix used for preparation of the calibration standards was not the authentic one. Nevertheless, the addition of SIL-IS corrected for the possible differences in the matrix effects between the calibrants and the samples. Therefore for the compound for which SIL-IS was added, similar concentrations were obtained with the standard addition and the external calibration approach. However, this was not the case for the compounds for which no SIL-IS. was added. This was considered to be due to the difference in the matrix effects between the calibrants and the samples. In the case of UDP-Glc the validation failed when using the standard addition approach due to the high background amounts of this compound in the spiking matrix. This was not the case with the external calibration since UDP-Glc was not present in the spiking matrix. Good linearity was obtained in the concentration range between 0.1-1.2 µg/ml.

An example of an application where matrix matched calibration was used for quantification is given in **Paper 2**. Since more than one sample needed to be quantified, the standard addition approach was not taken into account in this study. A mutant that contained small amounts of c-diGMP was obtained and therefore the extract of this mutant was used as the spiking matrix for preparing the calibrants. The advantage of having such a matrix allowed preparation of matrix matched calibration curve (R^2 of 0.99, Figure 8) without having saturation problems of the detector and linearity issues as in the case with UDP-Glc. Additionally, the results from the intraday precision experiments showed an RSD <15 % (n=3). SIL-IS for c-diGMP was not available. Therefore matrix matched calibrants that are processed through the sample

preparation were necessary in order to correct not only for the losses during the sample preparation procedure but also for the possible ion-suppression as a result of the matrix effects. Furthermore, the approach used for quantification is more straightforward when compared to the standard addition. The standard addition approach becomes more time consuming and requires more sample volume when multiple samples need to be quantified as in the case with *Ruegeria mobilis*.

However, it is not always possible to create a mutant that contains low background amounts of the compound of interest. Alternative to this could be the use of a fully ^{13}C labeled biomass to spike the standards for the calibration curve [61,80,113]. In that respect the standards will be prepared in the authentic matrix. However, in order to be able to use the commercially purchased ^{13}C SIL-IS e.g. [$^{13}C_{10}$]ATP that were already available in-house, an alternative approach for creating an authentic matrix was investigated (**Paper 5)**. The authentic matrix was prepared by growing *S. cerevisiae* in a medium that contained [$^{13}C_6$]-glucose/non labeled glucose (50/50, w/w) which, as expected, resulted in pools of metabolites with labeling in different carbon positions. The advantage of this matrix was that the pools of the compounds with only ^{12}C or ^{13}C carbons were very low or even not measurable and showed minimal or no interference to the spiked amount of non-labeled standards and their SIL-IS (Figure 16).

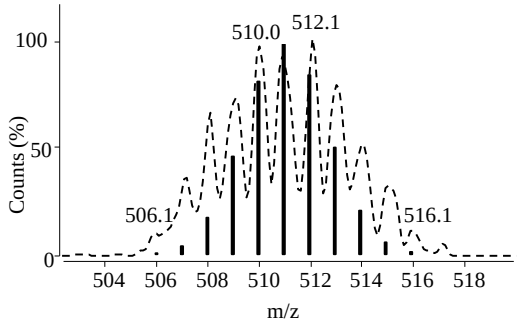


Figure 16. Measured (dashed line) and calculated (bars) isotopic patter of ATP extracted from *S. cerevisiae* cultivated in medium containing 50 % (w/w) [$^{13}C_6$]glucose/non-labelled glucose.

As a result of this, both the non-labeled and SIL-IS standards could be spiked in the matrix resulting in more straightforward validation and quantification. The highest interference from the matrix was to the MS signals of [$^{15}N_5$]ADP and [$^{15}N_5$]AMP while the lowest was to the coenzymes and the redox compounds (Figure 17).

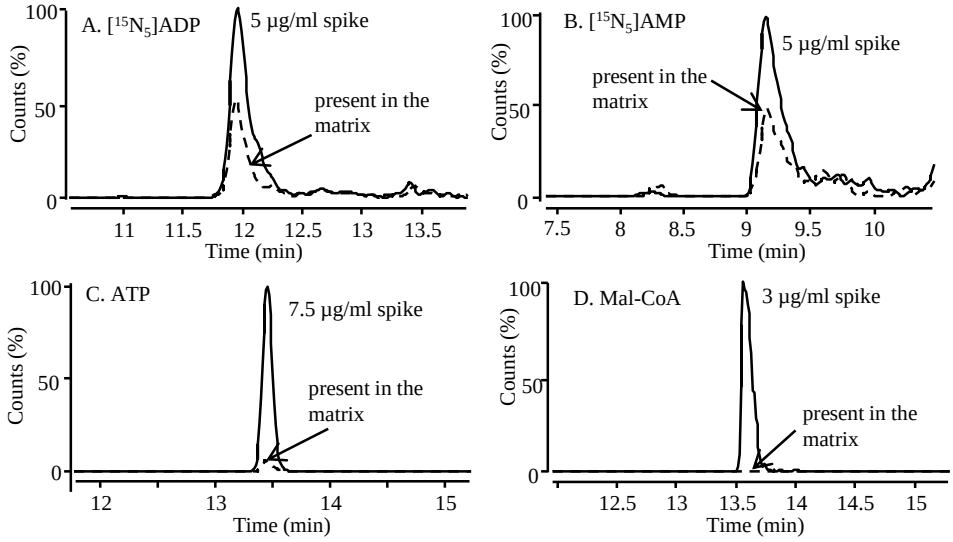


Figure 17. Superimposed chromatograms of [$^{15}N_5$]ADP, [$^{15}N_5$]AMP, ATP and Mal-CoA in the blank matrix (dashed line) and matrix spiked (solid line) with the corresponding standards with concentrations as given in the figure.

As expected, increase in the monoisotopic mass of 5 Da was observed for the ^{15}N labeled ADP and AMP when compared to the non-labeled compound standards. This increase was also detected in the biological matrix obtained by growing *S. cerevisiae* in media containing [$^{13}C_6$]glucose/non-labeled glucose (50/50, w/w) (Figure 10). Therefore choosing ^{13}C instead of ^{15}N labeled ADP and AMP would have resulted in less interference. However, [$^{13}C_{10}$]ADP and [$^{13}C_{10}$]AMP were not commercially available when the study was conducted.

The results from the validation showed good linearity over the inspected concentration range (R^2>0.99) as well as acceptable accuracy and precision (RSD ≤20 % and accuracy within 80-120 %) for all the compounds investigated. In the case of Ac-CoA, the difference in the monoisotopic mass between the non-labeled Ac-CoA and its SIL-IS was only 2 Da. This resulted in a cross signal contribution due to isotopic interference and increasing

concentration of the SIL-IS in the calibrants by increasing the concentration of the non-labeled Ac-CoA (Figure 18A).

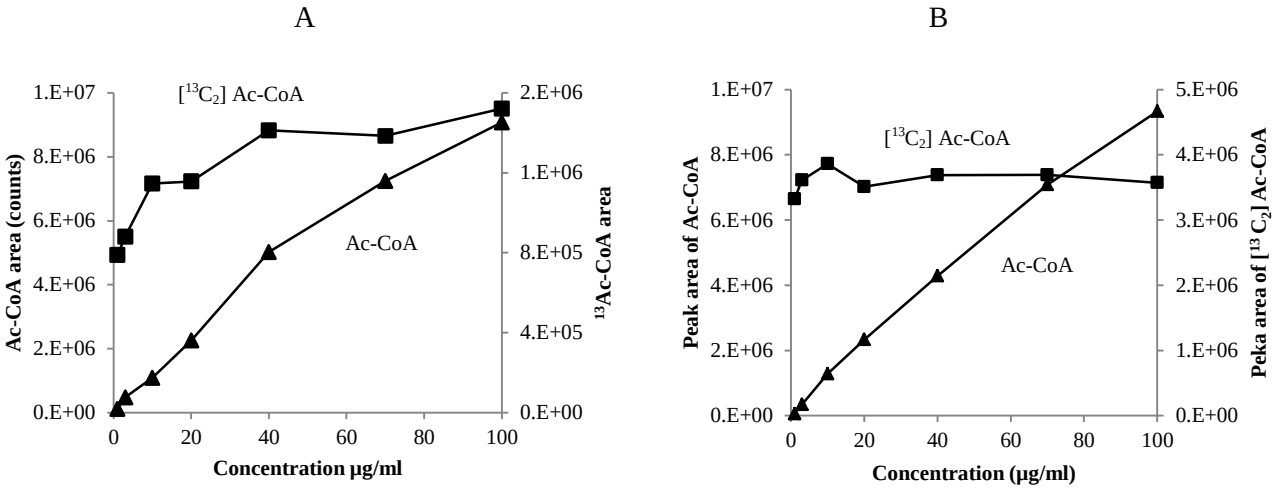


Figure 18. Response curves for A. Ac-CoA and 3 µg/ml [$^{13}C_2$] Ac-CoA and B. Ac-CoA and 10 µg/ml [$^{13}C_2$] Ac-CoA as a function of the Ac-CoA concentration.

The cross signal contribution caused significant non-linearity at high concentration levels, when low amounts of SIL-IS were used (Figure 19). However, the linearity was improved by increasing the amount of SIL-IS added to the calibrants [121] (Figure 19).

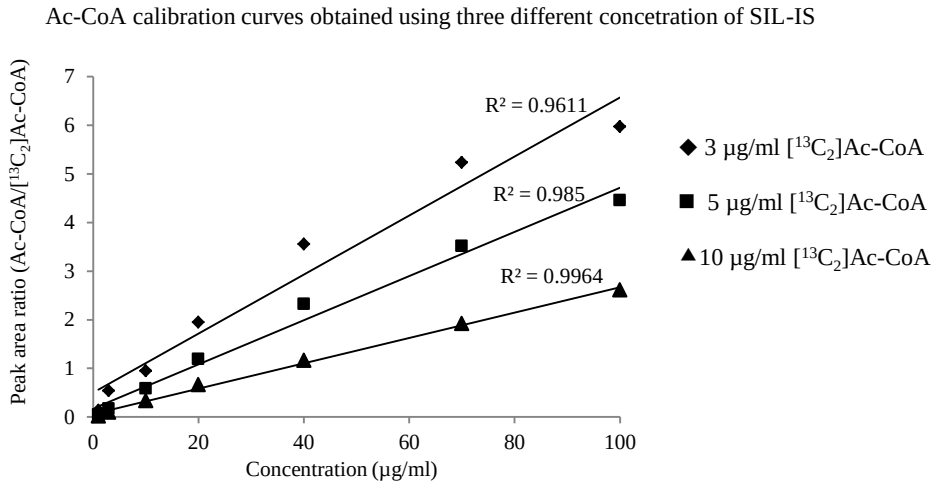


Figure 19. Calibration curves for Ac-CoA obtained using three different concentrations of the corresponding SIL-IS in this case [$^{13}C_2$]Ac-CoA.

Inspection of the matrix effects showed that the signal suppression at the beginning of the gradient (0-3 min) was related to the non-retained compounds that elute at the beginning of the analysis (Figure 20).

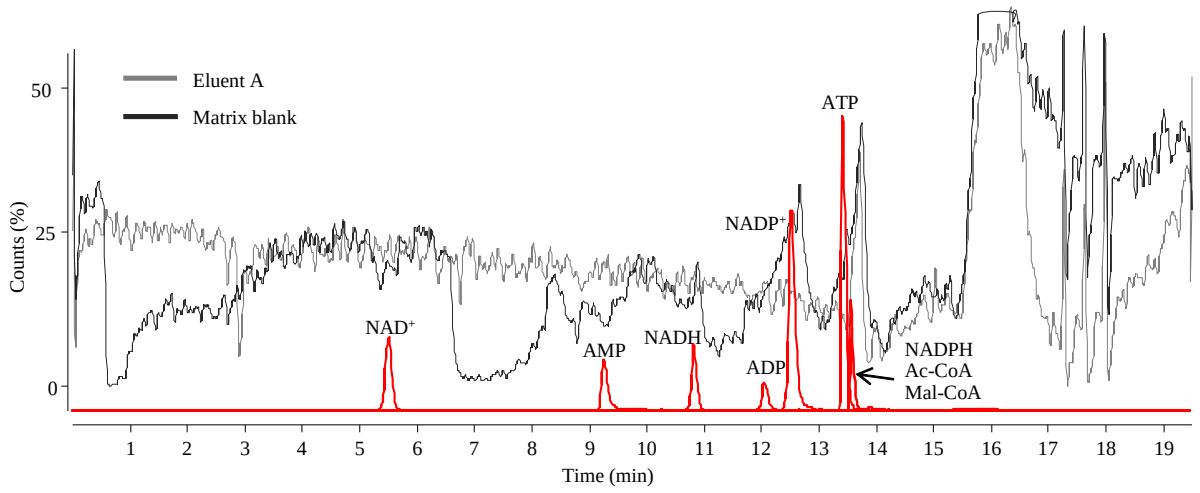


Figure 20. Overlaid chromatograms of post column infusion of 10 µg/ml ATP into the MS after injecting eluent A or matrix blank with superimposed chromatograms of all the analyzed compounds.

The suppression of the signal between 7 - 8 min. was due to the elution of PIPES, one of the buffering agents used during the extraction. Phosphate and EDTA affected the ion-signals of AMP and ADP, respectively. The phosphate was coming from the sample itself, while EDTA was used as a second buffering agent during the extraction. Furthermore, the validation showed that the SIL-IS successfully corrects for any matrix suppression and loss of analytes during the sample preparation, which indicates that a calibration curve could be prepared in the neat solution for the compounds for which SIL-IS is available.

Due to the absence of a matrix free of the analyte, the quantitation of the analytes in **Paper 3** and **Paper 4** was done using external calibration with standards prepared in the clean solvent. The main disadvantage of this approach is that the loss of the compounds during the sample preparation procedure and the matrix suppression will not be corrected unless a SIL-IS is added to both the samples and the calibration standards.

As previously mentioned in the introduction part, the validation and quantification approaches for the intracellular metabolites are less straightforward. Finding an appropriate approach and matrix for preparation of the calibration curves was one of the biggest challenges during this Ph.D. thesis. Furthermore, the inspection of the recoveries and the matrix effects was shown to be a very important step during the validation of the method in order to get reliable data especially for the compounds for which SIL-IS were not available. The use of SIL-IS showed to be a very powerful approach for i) overcoming the issues with losses of metabolites during the sample preparation method, ii) preventing the over or underestimation of metabolite levels due to matrix effects and iii) improving the overall performance of the method.

2.1.4 Summary of the findings using the targeted analysis with ion-pair LC-QqQ

This section summarizes the results obtained from the application of the developed ion-pair LC-QqQ method.

2.1.4.1 Measurement of the intracellular metabolites in *S. cerevisiae*

The two most commonly used *S. cerevisiae* strains as cell factories are: CEN.PK and S288C [122]. When these two strains were engineered to produce vanillin-β-glucoside (VG), difference in the VG level produced was observed. In order to investigate the possible reasons for this, various ‘’omics’’ tools were applied, among which measurement of a range of intracellular metabolites such as ATP, ADP, AMP, NAD^+, NADH, $NADP^+$, NADPH, UDP, UDP-Glc and UTP (Figure 21).

It should be noted that ATP, NADPH and UDP-Glc were directly involved in the production of VG. Three biological replicates for each strain were quenched and the intracellular metabolites were extracted with the most commonly used methods for quenching and extraction of yeast: -40°C MeOH/H_2O (60/40, v/v) and boiling EtOH/H_2O (75/25, v/v). The

high energy charge ratio of 0.91 obtained for all of the samples ensured good quenching of the metabolism and high quality metabolomics data.

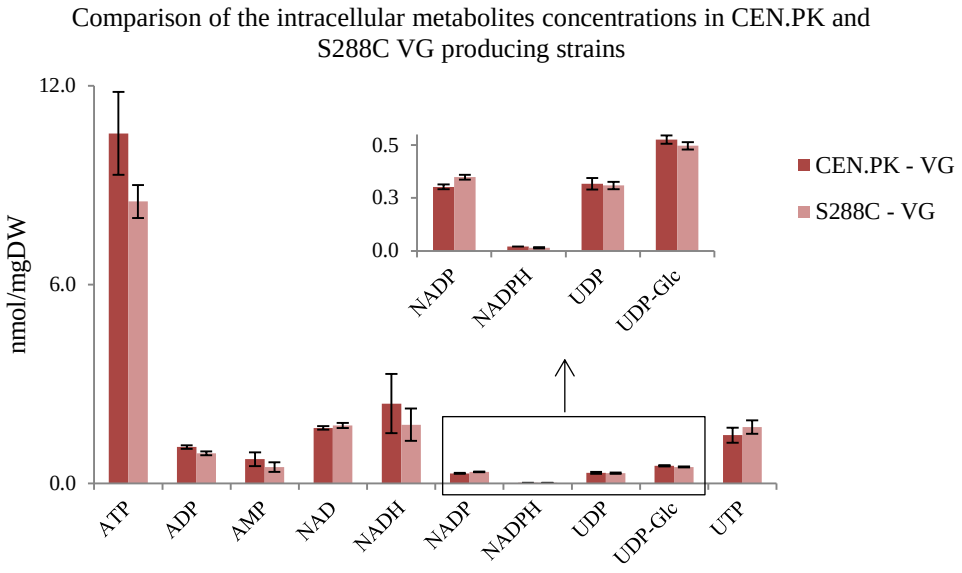


Figure 21. Diagram showing the measured intracellular concentrations of the metabolites extracted from CEN-PK and S288C vanillin-β-glucoside producing strains during steady state growth.

The concentrations of the intracellular metabolites between the VG producing CEN.PK (C-VG) and S288C (S-VG) strains appeared to be very similar. A slight difference was observed only for ATP (Figure 21). *Canelas et al.*[123] reported concentrations for ATP in the CEN.PK and S288C yeast strains which were not genetically engineered to produce VG. The ATP concentration ratio between the C-VG and S-VG strains in our study was found to be very similar to the one reported by *Canelas et al.* [123]. Therefore it was assumed that the differences in the ATP concentrations between C-VG and S-VG strains is not related to the differences in the VG production.

In the same study *Canelas et al.* [123] reported concentrations for ADP, AMP, UTP, UDP-Glc, NAD^+, $NADP^+$, and NADPH as well. When looking at the absolute values, the intracellular concentrations of ATP and AMP measured in the C-VG and S-VG were slightly higher than the ones reported in the literature, while ADP was slightly lower. UTP

concentrations were found to be very similar to the literature [123], while 5 times lower concentrations of UDP-Glc were measured in our study.

The slightly higher absolute values for ATP and AMP were speculated to be due to lab to lab differences. The slightly lower concentration of ADP measured in our study was speculated to be due to the absence of IS. Thus, the losses during the sample preparation, as well as the possible ionization suppression, were not corrected which might lead to underestimation of the concentration. The same was speculated for UDP-Glc. Therefore, inspection of the HR-MS spectra of the samples was performed. This was done, in order to investigate if there is another compound that elutes at the same time as ADP and UDP-Glc which might cause suppression and therefore underestimation of the concentration. At an RT where ADP elutes no other major ions were detected except for m/z 426.0221 that corresponds to ADP. Thus, the slightly lower concentration measured in our study might be due to losses during the sample preparation. However, at the RT where UDP-Glc elutes, high intensity peak with m/z of 96.96 was observed which corresponds to phosphate, thus causing strong suppression of the UDP-Glc MS signal. Regarding the concentrations of the redox compounds, *Canelas et al*. [123] reported only relative amounts and there was no clear explanation how these amounts were calculated. Therefore it was not possible to compare the values obtained for the redox compound in our study to those in the literature.

2.1.4.2 Measurement of nucleotides in *L. lactis*

The applicability of the developed ion-pair LC-QqQ method was assessed by measurement of ribo- and deoxyribonucleotides in *L. lactis*. In the current study 15 nucleotides were quantified including ATP, ADP and AMP which allowed determination of the ECR as a measure for the quality of the metabolomics data. In general, the intracellular pools of the ribonucleotides were higher than the deoxyribonucleotides. The deoxyribonucleotides were either close to the limit of detection or non-detectable. However, the concentrations of all

measured nucleotides using the ion-pair LC-QqQ method were higher than the concentration obtained using the ^{33}P labeling followed by thin layer chromatography (TLC) [92] (Figure 22).

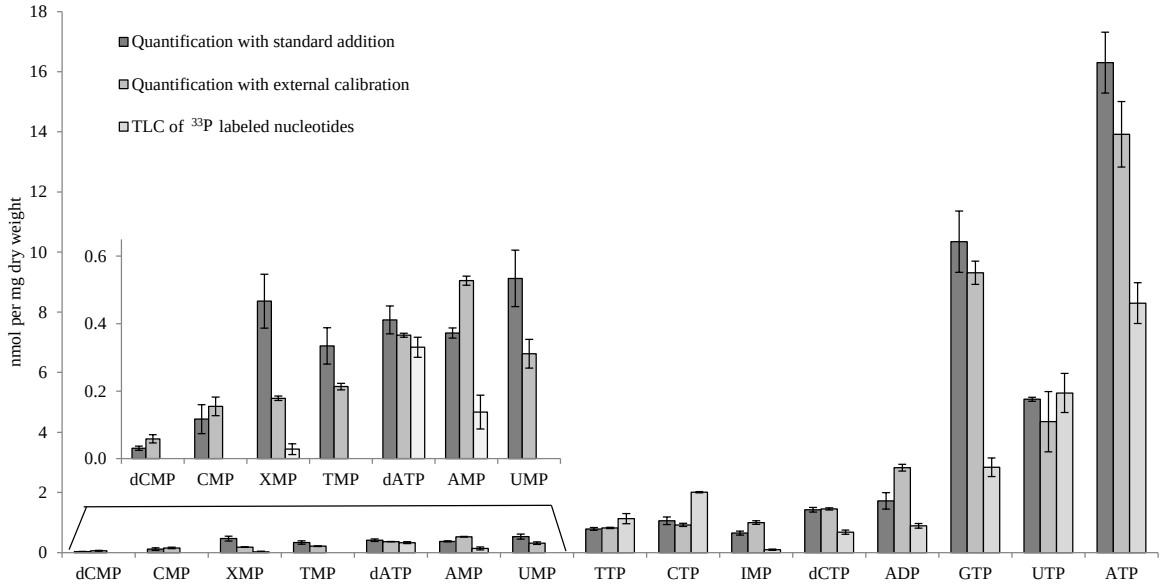


Figure 122. Comparison of the measured intracellular concentrations by ion-pair LCiMS/MS using two quantification approaches with the concentrations determined using ^{33}P labeling followed by TLC (dCMP, CMP, TMP, and UMP were not measured by TLC) [92].

Furthermore, the ratio between cytidine 5'-triphosphate (CTP) and deoxycytidine 5'-triphosphate (dCTP) has been found to be between 5 and 8 in the cells [124], however in this case slightly higher concentration of dCTP was measured giving a CTP/dCTP ratio close to 1. Even by growing *L. lactis* in a media that contained cytidine that should boost the CTP pool, the same ratio was obtained. On the contrary, the CTP/dCTP ratio in *S. cerevisiae* and CHO cells was found to have the expected value (between 5 and 8). In addition, similar concentration for CTP and dCTP were obtained with the two different quantification approaches (standard addition and external calibration). Therefore, the difference in the CTP/dCTP ratio between the current study and the literature was assumed to be due to physiology difference and not a technical error related to the analytical method.

2.1.4.3 Measurement of c-diGMP in *R. mobilis*

The main idea of this study (**Paper 2**) was to investigate whether the shift between motile and sessile life in the marine bacteria, *R. mobilis* F1926, is correlated to the intracellular concentration of c-diGMP. It has previously been shown that c-diGMP regulates the transition between planktonic and attached phenotype in other bacteria [125]. Therefore intracellular levels of c-diGMP were measured in: i) the wild type (WT) strain, ii) the pYedQ plasmid carrying mutant that contained a gene encoding c-diGMP synthesizing diguanylate cyclase, iii) the pYhjH plasmid carrying mutant that contained a gene encoding c-diGMP-degrading phosphodiesterase and iv) vector control strains F1926 pRK404A and F1926 pBBR1MCS3. It should be noted that for all of the strains, both shaken and static cultures were grown in order to mimic the motile and sessile lifestyle. In general for all the strains, the film forming static cultures contained higher amounts of c-diGMP than the shaken cultures, where no film was formed. Furthermore, the introduction of pYedQ plasmid into *Ruegeria mobilis* F1926 increased the levels of c-diGMP under both shaken and static conditions when compared to the other strains. In addition, more aggregates were formed in the pYedQ plasmid carrying mutants than in the WT in both shaken and stagnant cultures. The static cultures of pYhjH plasmid carrying mutant, showed decreased concentrations of c-diGMP when compared to the vector control strain. The addition of the pYhjH plasmid caused formation of more motile cells in the static cultures and prevented formation of aggregates in the static cultures. The data obtained showed that the developed IP-RP LC-QqQ method was capable of detecting the altered levels of c-diGMP in the mutants. Moreover, the data obtained, confirmed the hypothesis that c-diGMP is the key second messenger in the transition between motile and sessile life also in *R. mobilis.*

2.1.4.4 Measurement of the intracellular metabolites in CHO cells

The aim of this study was to investigate the effect on the central metabolism of an additional metabolic burden induced by recombinant protein production. Therefore seven clones with different EPO productivities were quenched followed by extraction of their metabolites. Different intracellular metabolites from the central carbon metabolism ranging from sugar phosphates, redox compounds as well as nucleotides were measured using the developed ion-pair LC-QqQ method. As described in **Paper 3** no differences in the ECR were detected among the CHO cells, showing that the cells energy metabolism is keeping up with the energy demand.

2.1.4.5 Measurement of intracellular metabolites in *M. corallina*

It should be mentioned that this was an EU project in collaboration with an italian company and due to confidentiality issues and lack of information, the data will be discussed only from analytical point of view.

The main idea of this project was to develop an economically viable production process for a lantibiotic which is produced by *M. corallina*. Two different strains were used: i) WT that was able to produce the lantibiotic and ii) a null strain that did not produce the lantibiotic. When growing on nitrate as a nitrogen source, the two strains showed difference in the production of the lantabiotic. In order to check if the differences in the lantibiotic production between the two strains results in any differences in the energy metabolism, IP-RP LC-QqQ was used to determine the intracellular pools of ATP, ADP, AMP and the redox pairs: NAD^+/NADH and $NADP^+$/NADPH. For this purpose both the WT and the null strain were grown in the same minimal medium with glucose as a carbon source in continuous cultivations and samples were taken from the chemostat when steady state was reached. The

samples were quenched using -40 °C MeOH/H_2O (60/40, v/v) and extracted using MeOH/CH_3Cl (2:1, v/v)

The results from the measurements of the intracellular metabolites are given in Table 7.

Table 7. Intracellular concentration of the measured metabolites in *M. corallina* expressed in nmol/mgDW.

	nmol/mgDW						
Strain	**ATP**	**ADP**	**AMP**	**NAD^+**	**NADH**	**$NADP^+$**	**NADPH**
WT	2.9 (1.4)[1]	1.7 (0.8)	0.2 (0.1)	1.4 (0.7)	ND	0.3 (0.1)	0.03 (0.009)
Null strain	1.5 (0.8)	0.5 (0.3)	0.1 (0.03)	0.7 (0.6)	0.05 (0.04)	0.2 (0.03)	0.03 (0.01)

[1]The numbers in the brackets are the standard deviations calculated from three technical replicates and two biological replicates.

The main problems encountered during this study were both organism and sample preparation method related. The organism dependent problems were related to the slow growth of this organism. The generation of the biomass was a long process (doubling time ~ 35 h), thus creating difficulties in the cases when some of the fermentations needed to be repeated. Consequently the frequency of sample generation was not as high as with the other organisms during this Ph.D. study, thus limiting the number of different sample preparation protocols that could be tested. The method dependent problems were related to the centrifugation as a technique for separation of the cells from the quenching supernatant. As explained in the sample preparation part, the pellet formed during the centrifugation was very easily re suspended in the quenching supernatant, thus resulting in losses of the biomass during the decanting. This was considered to be the reason for the big standard deviations observed between the technical (n=3) and the biological (n=2) replicates (Table 7). Longer time than 5 min for centrifugation was not considered in order to avoid exposure of the cells to the quenching supernatant and loss of the metabolites due to leakage. Furthermore,

quantification of the NAD/NADH and NADP/NADPH pairs was considered to be very challenging. As previously explained in the development of IP-RP part, oxidation of NADH and NADPH into NAD^+ and $NADP^+$, respectively was happening in the injection vial. This resulted in very high NAD^+/NADH and $NADP^+$/NADPH ratios. As previously explained re-dissolving the pure standards of NADH and NADPH in ammonium acetate pH 8, prevented the oxidation of these two compounds. Therefore re-dissolving the extracts into ammonium acetate pH 8 might solve the problem, however this needs to be further investigated.

2.1.4.6 Measurement of intracellular metabolites in *S. coelicolor*

The main idea behind the development of a method for quenching and extraction of intracellular metabolites from *S. coelicolor* was to investigate the effects of altered redox and energy levels in the metabolism on the production of the antibiotic actinorhodin (ACT). However, both the establishment of the sample preparation method (quenching and extraction) as well as the analysis of the redox compounds were challenging as explained previously.

ACT synthesis is a high energy demanding process, where 6 molecules of NADPH, 16 molecules of Ac-CoA and 16 molecules of ATP are utilized to produce one molecule of ACT. Consequently, it was expected that the production of ACT will be affected when the energy levels are perturbed. In this respect, two mutants were constructed by: i) overexpression of a *nox* gene that encodes an NADH oxidase that can oxidize NADH to NAD^+ (oxp-*nox*) and ii) overexpression of a *pos5* gene that encodes an NADH kinase that catalyzes the conversion of ATP and NADH into ADP and NADPH (oxp-*pos5*). The measurements of ACT in the two strains showed that the oxp-*nox* strain had an increased production of ACT while the oxp-*pos5* showed decreased ACT production. Since the perturbations made were expected to also have an impact on the energy metabolism, the concentrations of ATP, ADP, AMP and the redox compounds were measured in the WT, the

oxp-*nox* and oxp-*pos5* strains. The sampling was done at 6 different TP during batch fermentation: i) 3 TP during the exponential and ii) 3 TP during the antibiotic producing phase. It should be mentioned that the data regarding the redox compounds were not processed at the time when the thesis was written due to the lack of time. Therefore, only the concentrations of ATP, ADP and AMP are presented in this thesis.

As explained previously in the part regarding the establishment of the quenching and extraction procedures, leakage was observed when *S. coelicolor* was quenched with the cold MeOH. Figure 23 shows the measured concentrations of the ATP, ADP and AMP in the biomass and the supernatant of the WT and the *oxp-nox* mutant. It should also be noted that due to lack of time the supernatants from the oxp-*pos5* strain were not processed and therefore data on the leakage is presented only for the WT and the oxp-*nox.* Although leakage was observed, when looking at Figure 23, trend could be seen for the measured concentrations of ATP, ADP and AMP in the WT and the *oxp-nox* mutant.

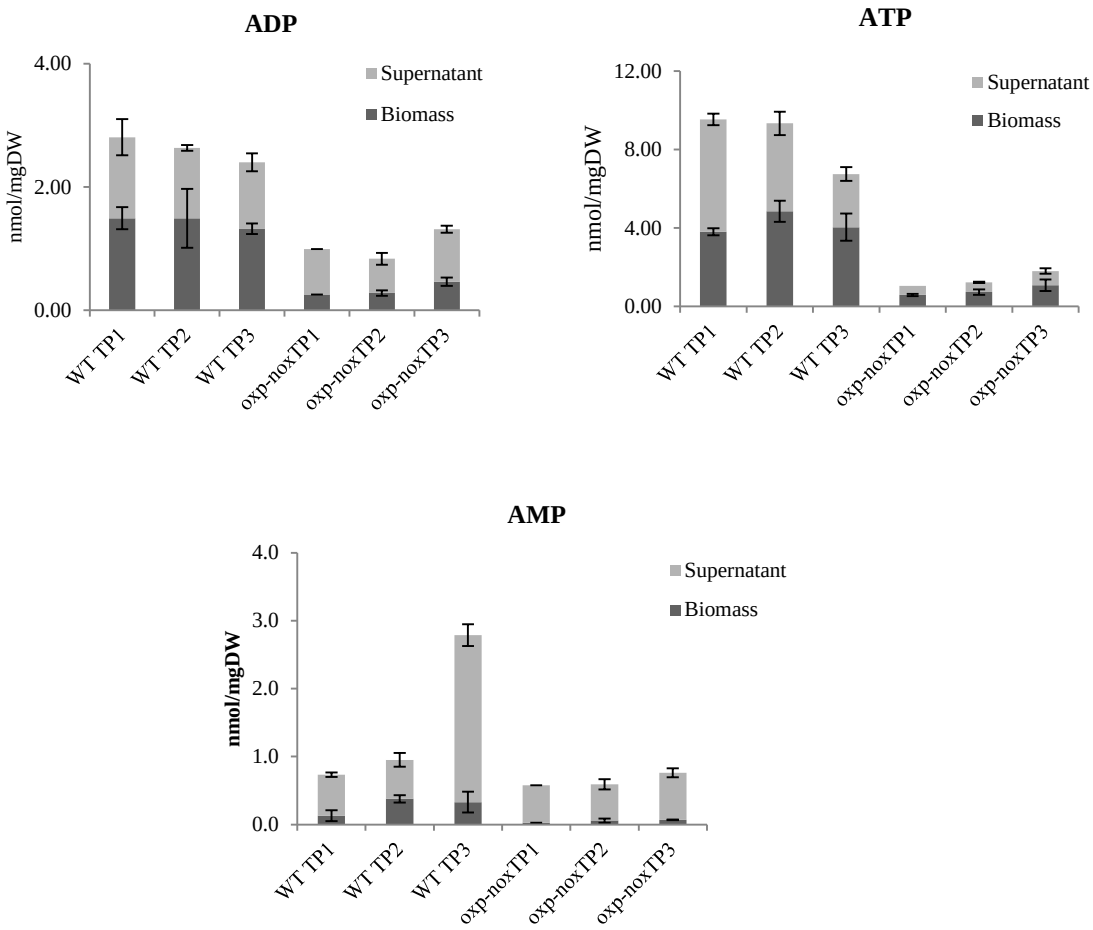


Figure 23. Concentrations of ATP, ADP and AMP measured intracellular and in the quenching supernatant at 3 TP in exponential phase from the WT and the oxp-*nox* strain of *S. coelicolor*. The error bars indicate standard deviations from three technical replicates.

Furthermore, the data from the measured concentrations of ATP, ADP and AMP in the biomass form the three strains (WT, *oxp-nox*, *oxp-pos5*) at 3 TP (TP1-3) during the exponential and 3 TP during the antibiotic production phase are given in Table 8.

Table 8. Concentrations of ATP, ADP and AMP in the WT, oxp-*nox* and the oxp-*pos5*, measured in the biomass samples at 3 different TP during the exponential phase and 3TP during the antibiotic production phase.

Sampling phase	Strain	nmol/mgDW			ECR[1]
		ATP	ADP	AMP	
Exponential phase	WT TP1	3.82 (0.18)[2]	1.49 (0.18)	0.13 (0.08)	0.84
	WT TP2	4.86 (0.54)	1.65 (0.43)	0.35 (0.07)	0.83
	WT TP3	4.04 (0.69)	1.32 (0.08)	0.33 (0.15)	0.83
	oxp-*nox* TP1	0.59 (0.05)	0.24(0.02)	0.02 (0.003)	0.83
	oxp-*nox* TP2	0.74 (0.14)	0.28 (0.04)	0.06 (0.03)	0.82
	oxp-*nox* TP3	1.09 (0.30)	0.46 (0.07)	0.07 (0.005)	0.82
	oxp-*pos5* TP1	2.04 (0.80)	0.83 (0.15)	0.68 (0.02)	0.69
	oxp-*pos*5TP2	1.17 (0.29)	0.49 (0.12)	0.41 (0.04)	0.69
	oxp-*pos5* TP3	1.04 (0.33)	0.39 (0.16)	0.31 (0.05)	0.71
Antibiotic production phase	WT TP1	4.34(0.14)	0.37 (0.02)	0.38 (0.02)	0.89
	WT TP2	3.24 (0.08)	0.26 (0.02)	0.30 (0.01)	0.89
	WT TP3	1.83 (0.04)	0.19 (0.03)	0.23 (0.04)	0.85
	oxp-*nox* TP1	1.66 (0.16)	0.79 (0.09)	1.40 (0.08)	0.53
	oxp-*nox* TP2	1.89 (0.28)	1.22 (0.71)	1.82 (0.17)	0.51
	oxp-*nox* TP3	1.32 (0.34)	1.19 (0.36)	2.06 (0.63)	0.42
	oxp-*pos5* TP1	1.88	1.09	0.92	0.62
	oxp-*pos5* TP2	1.66	1.26	0.99	0.59
	oxp-*pos5* TP3	2.08 (0.02)	1.20 (0.16)	1.21 (0.02)	0.60

[1]ECR- Energy charge ratio calculated as $(ATP + 0.5ADP)/(ATP + ADP + AMP)$

[2]The numbers in the brackets are the standard deviations calculated from three technical replicates.

[3]No standard deviation given due to lack of replicates

In the exponential phase, the *oxp-nox* strain showed lower ATP concentrations when compared to the WT. This was due to an increased oxidation of NADH and consequently lower production of ATP during oxidative phosphorylation. In the antibiotic production phase there was an increase of the ATP concentration due to the decreased growth.

In comparison to the oxp-*nox* mutant, increased concentration of ATP was detected in the oxp-*pos5* mutant and was speculated to be due the higher NADPH production. In order to maintain the NADPH/NADP ratio, higher ATP is produced. However, we could not explain why the ATP concentration decreased during the antibiotic phase of the oxp-*pos5* strain.

Crucial part for this study will be the comparison of the data for $NADH/NAD^+$ and $NADPH/NADP^+$ pools between the mutant and the reference strains. As previously explained

in the thesis, problem with the oxidation of NADH and NADPH was observed. Since re-dissolving the standards of these compounds into ammonium acetate pH 8 decreased the oxidation, re-dissolving the extract into the same solution might be taken into consideration in order to inspect if this will reduce the oxidation in the samples as well. However, in order to get reliable data and confirm the assumptions presented here, quenching method that will result in less leakage needs to be established.

2.2 Multitargeted approach using IP-RP LC-Q-TOF MS

The main idea of the coupling the ion-pair method to the Q-TOF-MS instrument was to explore the applicability of the high resolution instrument for simultaneous identification of as many intracellular metabolites as possible (**Paper 6**). The feasibility of using the Q-TOF was evaluated with identification of compounds from three different organisms: *S. cerevisiae*, *M. corallina* and *S. coelicolor*.

The multitargeted approach in this study included: i) aggressive dereplication of the full scan HR-MS data using search lists of known compounds and ii) MS/HRMS data searched in METLIN library. Both auto and all ions MS/MS approaches were investigated for obtaining fragmentation spectra. As previously explained in the introduction part, the auto MS/MS approach was based on the selection of a precursor ion by the quadrupole that is further on fragmented in the collision cell, while in the all ions MS/MS (MS^E) approach all the ions generated are fragmented using low and high CE values. However, the all ions MS/MS approach was not suitable for the current study due to the similarities in the fragmentation patterns between the intracellular metabolites. For example all phosphorylated metabolites gave m/z 96.96 and m/z 78.959 and were primarily chosen by the software as qualifier ions. These two ions were considered to not be specific enough to confirm the identity of the compounds. Therefore, auto MS/MS approach was used for acquiring the MS/HRMS data.

The principle of the screening method was firstly to analyze the sample extract from the organism of interest using the ion-pair Q-TOF MS. Aggressive dereplication was done as described in *Kildegaard et al.* [126] by creating *.csv files that contained the empirical formula and the name of the compounds listed in an in-house database created in the ACD chemfolder format. This database was created by importing the online available metabolom databases of *Escherichia coli* (*E. coli* metabolome database, http://www.ecmdb.ca/) and yeast (Yeast Metabolome Database, http://www.ymdb.ca/) in the sdf format. Tentative identification was based on matching the accurate mass and isotopic pattern of the compounds in the search lists to that measured by the instrument. If a compound from the search lists was detected in the samples, the chromatographic peak that corresponds to that compound was colored. The aggressive dereplication was shown to be a fast approach for tentative identification of hundreds of compounds by using their accurate mass and isotopic pattern. The search lists could easily be changed and made more specific, which was an advantage when screening for only one group of compounds (e.g. nucleotides) was needed. Furthermore, the aggressive dereplication using search lists provided an ability to quickly get an overview of the possible isomers assigned to one peak.

Although the data analysis approach described in this study yielded putative identifications for several hundreds of compounds in a very short time, one should also be aware of the pitfalls of this approach. Very often one peak was assigned to more than one compound and vice versa. If the adduct of one compound had the same mass as the molecular ion of another compound, the software was reporting two hits for the same peak. Therefore a manual inspection of the results was required which was considered as a bottleneck of the approach. Even though the approach was helpful in giving a quick overview of the structural isomers present in the sample, it was not able to distinguish between them unless an RT was available. Furthermore, an in-source fragmentation was also shown to lead to miss-

identification. This was the case when an in-source fragment of one compound had the same elemental composition as the molecular ion of another compound, as shown for ADP-ribose 1″ -2″ cyclic phosphate and $NADP^+$, where an in-source fragment of $NADP^+$ had the same mass as ADP-ribose 1″ -2″ cyclic phosphate. In this respect the RT would be beneficial for identification, however a standard of ADP-ribose 1″ -2″ cyclic phosphate was not available in-house.

To overcome the above mentioned issues HRMS/MS dereplication was performed. This dereplication procedure was based on usage of HRMS/MS library for screening of the data acquired in auto MS/MS mode. The library allows the user to compare both the accurate mass and fragmentation data acquired form the sample with those from the library. The spectral library was created by analyzing the standards available in-house at three different collision energies (10, 20 and 40 eV) due to the different energy needed to fragment the different compounds. It should be noted that the creation of the library itself is a labor intense work since all the standards need to be analyzed using the IP-RP LC-QTOF MS methods and their fragmentation patterns are manually added into the library. However due to the fact that only limited number of primary metabolite standards were available in-house, the online METLIN library was mainly used. The HRMS/MS dereplication approach improved the efficiency and the confidence of the identification results. However, in the case of structural isomers that generated similar fragmentation patterns, (e.g. sugar phosphates) RT was the key parameter to differentiate between them. The limiting factor of the HRMS/MS dereplication was the size of the library and therefore when an MS/MS spectrum for a certain compound was not available in METLIN, manual inspections of the fragmentations spectra was performed for losses of e.g. phosphate, carboxylic groups etc.

Figure 24 shows the chromatograms of some of the identified intracellular metabolites in *S. cerevisiae*, *M. corallina* and *S. coelicolor*.

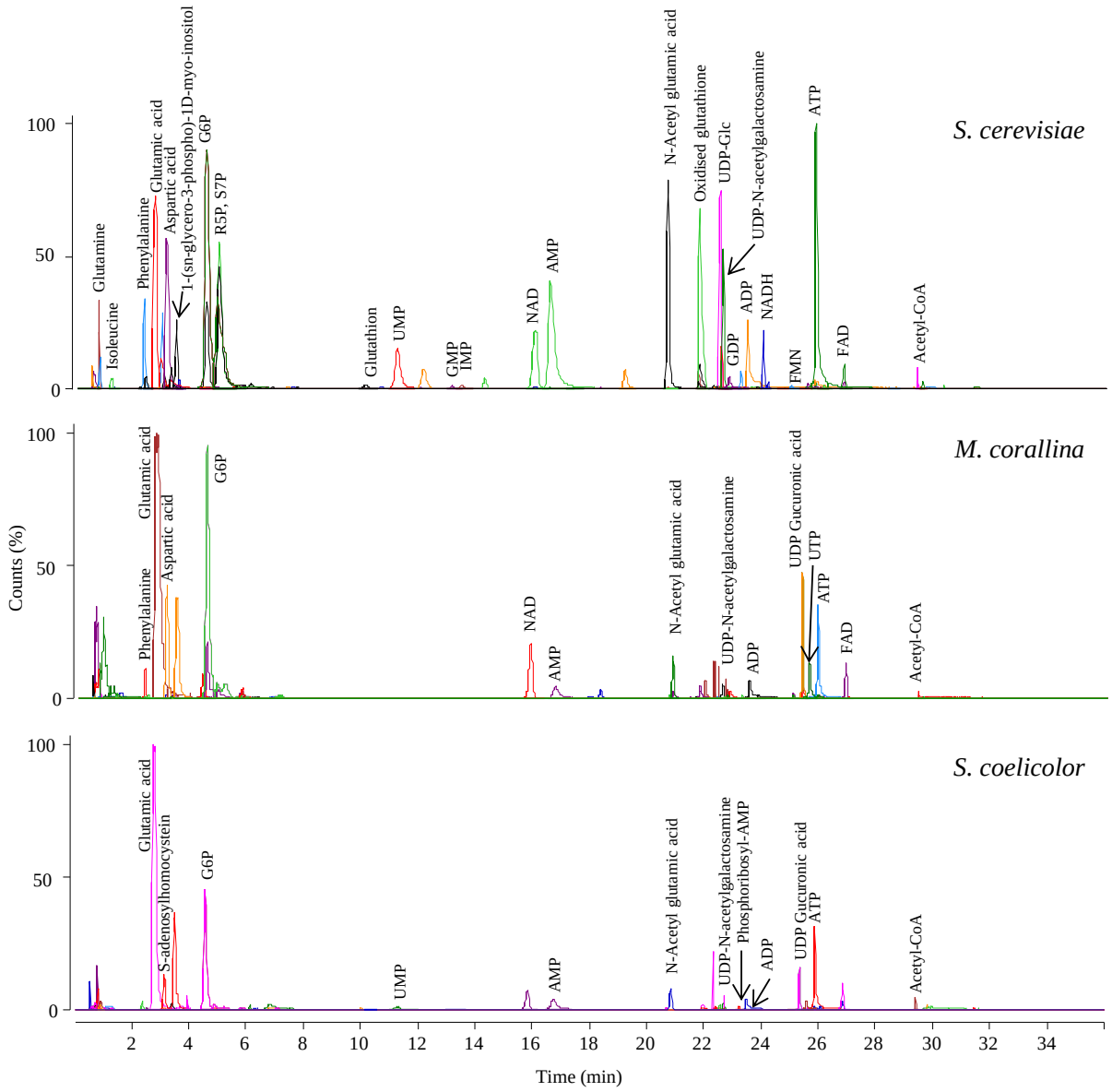


Figure 24. Extracted ion chromatograms of the most intense peaks identified in *S. cerevisiae*, *M. corallina* and *S. coelicolor* using the aggressive dereplication in combination with the MS/HRMS data searched in METLIN library.

In general the coupling of the ion-pair chromatography with Q-TOF-MS was shown to provide an opportunity to detect and identify many compounds. Having the full scan HR-MS data is very beneficial since they allow retrospective data analysis and search for other interesting target compounds that have been measured but have not been of interest for the current study. Processing the data using statistical tools can help in finding metabolites which

concentrations have been changed due to genetic manipulations or stress conditions. Furthermore, having the HR-MS and MS/HRMS data identification of the compounds can often be achieved without the need of standard compounds. The full scan data can help to identify the causes for suppression/enhancement of the MS signal. This information may help in improvement of the sample preparation method.

3. Future perspectives

Despite the great progress within the microbial metabolomics, further research, especially on the sample preparation techniques, is highly required. This will help in understanding which sample preparation methods should be used for which type of organisms and analysis. Furthermore, knowing the shortcomings of the methods applied will help in understanding the limitations of the data collected. To minimize the variability between the laboratories and analysts, more standardized methods are highly needed.

In order to improve the quenching of the bacteria the alternatives could include: i) measuring the intracellular metabolites in the total broth using the cold methanol quenching or ii) fast filtration since it could also solve the problem with instability of the biomass pellet of *M. corallina* obtained after centrifugation which resulted in biomass losses during decanting. Furthermore, standardized sampling system is highly required for future analysis of intracellular metabolites. This will increase the sampling frequency and will decrease the variations in the sample volume taken.

For improving the analysis, lowering the amount of buffers used during the sample preparation would be very beneficial, especially in the MeOH/chloroform extraction. In the case when organic acids need to be analyzed a selective removal of phosphate using anion exchange SPE and low pH could be considered as a solution [105].

Ammonium acetate (pH 8) showed to reduce the oxidation of NADH to NAD^+ and NADPH to $NADP^+$. By processing the standards prepared in the ammonium acetate through the sample preparation, one can investigate if oxidation occurs during the sample handling. Furthermore, re-dissolving of the sample extracts into ammonium acetate (pH 8) instead of TBA might improve the analysis of these compounds. However, further investigation of the long term stability of these compounds in the ammonium acetate solution is highly required.

Moreover, standardized analytical method validation strategies and reference materials for intracellular metabolites are sorely missing. These strategies are highly needed in order to confirm the accuracy and validity of the quantitative data acquired. Therefore, a debate in order to define the conditions and the metabolites for which this material will be produced is necessary in future. In respect to the validation of the analytical methods, the approach reported by *Mashego et al*. [113] which is an alternative to the SIL-IS approach used during this thesis, could be taken into consideration for future studies. In this approach extracts form cells grown on ^{13}C labeled glucose are used as SIL-IS. The advantage of the *Mashego et al.* [113] approach is that SIL-IS will be obtained for compound for which there are no commercially available SIL-IS, thus improving the reliability of the quantitative data for those compounds.

Finally, as proposed by *Nielsen* and *Oliver* [127] and Griffin [128], building a database that will contain accurately measured metabolites under standardized conditions can serve as a good reference and a good assessment of the developed methods used for quantifying intracellular metabolites.

4. Conclusion

An IP-RP LC-MS based method was established for targeted and multitargeted analysis of intracellular metabolites from various microorganisms.

The optimization of the detection resulted in optimized MRM transitions for more than 50 metabolites including nucleotides, coenzymes, sugar phosphates and organic acids. IP-PR chromatography showed to give a good compromise between the retention and the separation of the metabolites of interest. The IP-RP method allowed separation of compounds with the same elemental composition (e.g. sugar phosphates, AMP/dGMP, G3P/DHAP) which was an advantage due to their similar fragmentation pattern.

Several different quenching and extraction methods were tested for yeast, bacteria, filamentous fungi and mammalian cells. High ECR values (0.80-0.95) were obtained for yeast when cold MeOH was combined with either boiling EtOH or MeOH/chloroform as extraction methods. Formic acid showed to be a good quenching method for *L. lactis* but not for the filamentous bacteria *M. corallina* and *S. coelicolor*. The reason for this was speculated to be morphology related. Several other quenching techniques were tested for the filamentous bacteria based on saline, glycerol or cold methanol. The main problems observed were either low ECR or leakage during the quenching. As in the case for the filamentous bacteria, leakage during the cold methanol quenching was also observed for filamentous fungi. Ice cold 0.9 % (w/v) NaCl combined with MeOH/ACN as an extraction was shown to be able to stop the mammalian cell metabolism (ECR of 0.92). In this study, metabolites from glycolysis, TCA and PPP pathway were analyzed.

The sample preparation showed to be very important step not only for obtaining meaningful metabolomics data but also for the analysis step itself. The buffering agents (e.g. EDTA or PIPES) used in the sample preparation as well as the highly abundant compounds present in the cells, such as phosphate, were shown to cause suppression of the LC-MS signal.

The addition of SIL-IS after the quenching was shown to be very important for i) correction of metabolites losses during the sample preparation, ii) correction of suppression caused by coeluting compounds and iii) improvement of the linearity. It was shown that for good linearity the amount of the SIL-IS should be in sufficient amount, either close to the amount of non-labeled metabolite that need to be quantified or close to the middle concentration point on the calibration curve. The validation approaches used along this study shown acceptable accuracy/recovery (80-120 %) and precision (RSD $\leq$ 20 %) as well as good linearity confirming the reliability of the data obtained.

An approach for creating a spiking matrix based on ^{13}C labeling was established as well. This allowed preparation of the quality control samples in the matrix, by spiking both SIL-IS and non-labeled standards. One of the requirement here is that the ^{13}C SIL-IS should be used instead of ^{15}N labeled in order to minimize the matrix interferences.

Finally, coupling of the IP-RP to a Q-TOF showed to be a powerful tool for identification of known unknowns without the need of a standard. The analysis resulted in identification of 60 compounds from different microorganisms such as *S. coelicolor*, *M. corallina* and *S. cerevisiae*. The identification was based on their HR-MS and MS/HRMS spectra.

The data presented in this thesis, showed that the established analytical and sample preparation methods are a valuable addition to the ‘’omics tools’’ use to reveal key information regarding the biochemical processes within the cells.

5. References

[1] S. G. Villas-Bôas, S. Mas, M. Åkesson, J. Smedsgaard, J. Nielsen, Mass spectrometry in metabolome analysis, Mass Spectrom. Rev. 24 (2005) 613-646.

[2] O. Fiehn, Combining genomics, metabolome analysis and biochemical modeling to understand metabolic networks, Comp. Funct. Genom. 2 (2001) 155-168.

[3] O. Fiehn, Metabolomics – the link between genotypes and phenotypes, Plant. Mol. Biol. 48 (2002) 155-171.

[4] K. Dettmer, P. A. Aronov, B. D. Hammock, Mass spectrometry-based metabolomics, Mass Spectrom. Rev. 26 (2007) 51-58.

[5] J. W. Allwood, D. I. Ellis, R. Goodacre, Metabolomic technologies and their application to the study of plants and plant-host interactions, Physiol. Plantarum 132 (2008) 117-135.

[6] J. Föster, I. Famili, P. Fu, P. Ø. Palsson, J. Nielsen, Genome-scale reconstruction of *Saccharomyces cerevisiae* metabolic network, Genome Res. 13 (2003) 244-253.

[7] M. J Herrgård, N. Swainston, P. Dobson, W. B Dunn, K. Y. Arga, M. Arvas, N. Blüthgen, S. Borger, R. Costenoble, M Heinemann, M. Hucka, N. Le Novère, P. Li, W. Liebermeister, M. L Mo, A. P. Oliveira, D. Petranovic, S. Pettifer, E. Simeonidis, K. Smallbone, I. Spasié, D. Weichart, R. Brent, D. S Broomhead, H. V Westerhoff, B. Kürdar, M. Penttilä, E. Klipp, B. Ø Palsson, U. Sauer, S. G Oliver, P. Mendes, J. Nielsen, D. B Kell, A consensus yeast metabolic network reconstruction obtained from a community approach to system biology, Nat. Biotechnol. 26 (2008) 1155-1160.

[8] W. Weckwerth, Metabolomics in systems biology, Annu. Rev. Plant. Biol. 54 (2003) 669-689.

[9] E. Nevoigt, Progress in Metabolic Engineering of *Saccharomyces cerevisiae*, Microbiol. Mol. Biol. Rev. 72 (2008) 379-412.

[10] J. H. Park, S. Y. Lee, T. Y. Kim, H. U. Kim, Application of systems biology for bioprocess development, Trends. Biotechnol. 26 (2008) 404-412.

[11] W. C. Ruder, T. Lu, J. J. Collins, Synthetic biology moving into the clinic, Science 333 (2011) 1248–1252.

[12] J. L. Adrio, A. L Demain, Recombinant organisms for production of industrial products, Bioeng. Bugs. 2 (2010) 116–131.

[13] I. K. Kim, A. Roldão, V. Siewers, J. Nielsen, A systems-level approach for metabolic engineering of yeast cell factories, FEMS Yeast Res. 12 (2012) 228–248.

[14] J. L. Ingraham, O. Maaløe, F. C. Neidhardt, Growth of the bacterial cell, Sinauer Associates, Inc, Sunderland, MA, 1983, p. 87–173.

[15] F. C. Neidhardt, J. L Ingraham, M. Schaechter, Physiology of the bacterial cell. A molecular approach, Sinauer Associates, Inc, Sunderland, MA, 1992, p. 133–173.

[16] S. Sudarsan, S. Dethlefsen, L. M. Blank, M. S. Herzberg, A. Schmida, The functional structure of central carbon metabolism in *Pseudomonas putida* KT2440, Appl. Environ. Microbiol. 80 (2014) 5292–5303.

[17] B. Luo, K. Groenke, R. Takors, C. Wandrey, M. Oldiges, Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway and tricarboxylic acid cycle by liquid chromatography–mass spectrometry, J. Chromatogr. A, 1147 (2007) 153–164.

[18] M. Gomelskyc, AMP, c-di-GMP, c-di-AMP and now cGMP: Bacteria use them all!, Mol. Microbiol. 79 (2011) 562–565.

[19] J. I. Wheeler, L. Freihat, H. R. Irving, A cyclic nucleotide sensitive promoter reporter system suitable for bacteria and plant cells, *BMC Biotechnol.* 2013 (13) 1-10.

[20] R. M. Corrigan, A. Gründling, Cyclic di-AMP: another second messenger enters the fray, Nat. Rev. Microbiol. 11 (2013) 513–524.

[21] K. A. Mc Donough, A. Rodriguez, The myriad roles of cyclic AMP in microbial pathogens: From signal to sword, Nat. Rev. Microbiol. 10 (2012) 27–38.

[22] M. R. Mashego, K. Rumbold, M. De Mey, E. Vandamme, W. Soetaert, J. J. Heijnen, Microbial metabolomics: past, present and future methodologies, Biotechnol. Lett. 29 (2007) 1-16.

[23] M. Sauer, D. Mattanovich, Construction of microbial cell factories for industrial bioprocesses, J. Chem. Technol. Biot. 87 (2012) 445-450.

[24] S. Dietmair, N. E. Timmins, P. P. Gray, L. K. Nielsen, J. O. Krömer, Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol, Anal. Biochem. 404 (2010) 155–164.

[25] S.G. Villas-Bôas, U. Roessner, M.A.E. Hansen, J. Smedsgaard, J. Nielsen (Eds.), Metabolome Analysis: An introduction, Wiley–Interscience, Hoboken, NJ, 2007, pp. 39–82.

[26] W. de Koning, K. van Dam, A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH, Anal. Biochem. 204 (1992) 118–123.

[27] S.G. Villas-Bôas, J. Højer-Pedersen, M. Åkesson, J. Smedsgaard, J. Nielsen, Global metabolite analysis of yeasts: Evaluation of sample preparation methods, Yeast 22 (2005) 1155–1169.

[28] W. M. van Gulik, Fast sampling for quantitative microbial metabolomics Curr. Opin. Biotechnol. 21 (2010) 27–34.

[29] M. Rizzi, M. Baltes, U. Theobald, M. Reuss, In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model, Biotechnol. Bioeng. 55 (1997) 592–608.

[30] Y. Ohashi, A. Hirayama, T. Ishikawa, S. Nakamura, K. Shimizu, Y. Ueno, M. Tomitaab, T. Soga, Depiction of metabolome changes in histidine-starved *Escherichia coli* by CE-TOFMS, Mol. Bio Syst. 4 (2008) 135–147.

[31] G. J. Patti, O. Yanes, Gary Siuzdak, Metabolomics: the apogee of the omics trilogy Nat. Rev. Mol. Cell. Biol. 13 (2012) 263-269.

[32] E. Dudley, M. Yousef, Y. Wang, W. J. Griffiths, Targeted metabolomics and mass spectrometry, Adv. Protein Chem. Struct. Biol. 80 (2010) 45–83.

[33] L.D. Roberts, A. L. Souza, R. E. Gerszten, C. B. Clish, Targeted Metabolomics, Curr. Protoc. Mol. Biol. 98 (2012) Chapter 30: Unit 30.2.1–30.2.24.

[34] J. Hou, N. F. Lages, M. Oldiges, G. N. Vemuri, Metabolic impact of redox cofactor perturbations in *Saccharomyces cerevisiae*, Metab. Eng. 11 (2009) 253–261.

[35] D. Visser, G. A. van Zuylen, J. C. van Dam, M. R. Eman, A. Pröll, C. Ras, L. Wu, W. M. van Gulik, J. J. Heijnen, Analysis of in vivo kinetics of glycolysis in aerobic *Saccharomyces cerevisiae* by application of glucose and ethanol pulses, Biotechnol. Bioeng. 88 (2004) 157-167.

[36] A. B. Canelas, W. M van Gulik, J. J. Heijnen (2008) Determination of the cytosolic free NAD/NADH ratio in *Saccharomyces cerevisiae* under steady-state and highly dynamic conditions. Biotechnol. Bioeng. 100 (2008) 734–743.

[37] V.M. Boer, C. A. Crutchfield, P. H. Bradley, D. Botstein, J. D. Rabinowitz, Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations, Mol. Biol. Cell 21 (2010) 198-211.

[38] A. Wentzel, H. Sletta, S. Consortium, T. E. Ellingsen and P. Bruheim, Intracellular metabolite pool changes in response to nutrient depletion induced metabolic switching in S*treptomyces coelicolor*, Metabolites 2 (2012) 178-194.

[39] D.Y. Lee, B. P. Bowen, Trent R. Northen, Mass spectrometry-based metabolomics, analysis of metabolite-protein interactions and imaging, BioTechniques 49 (2010) 557-65.

[40] C. A. Smith, G. O'Maille, E. J. Want, C. Qin, S. A. Trauger, T. R. Brandon, D. E. Custodio, R. Abagyan, G. Siuzdak, METLIN: a metabolite mass spectral database, Ther. Drug. Monit. 27 (2005) 747-51.

[41] C. Wagner, M. Sefkow, J. Kopka, Construction and application of a mass spectral and retention time index database generated from plant GC/EITOF-MS metabolite profiles, Phytochemistry 62 (2003) 887–900.

[42] A. Ruiz-Aracama, A. Peijnenburg, J. Kleinjans, D. Jennen, J. van Delft, C. Hellfrisch, A. Lommen, An untargeted multi-technique metabolomics approach to studying

intracellular metabolites of HepG2 cells exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin, BMC Genomics, 12 (2011) 2-19.

[43] W. Lu, B. D. Bennett, J. D. Rabinowitz, Analytical strategies for LC–MS-based targeted metabolomics, J. Chromatogr. B 871 (2008) 236–242.

[44] J. Martin Büscher, D. Czernik, J. C. Ewald, U. Sauer, N. Zamboni, Cross-Platform comparison of methods for quantitative metabolomics of primary metabolism, Anal. Chem. 81 (2009) 2135–2143.

[45] B. Zhou, J. F. Xiao, L. Tuli, H. W. Ressom, LC-MS-based metabolomics, Mol. Bio Syst. 8 (2012) 470–481.

[46] A. Porzel, M. A. Farag, J. Mülbradt, Ludger A. Wessjohann, Metabolite profiling and fingerprinting of Hypericum species: a comparison of MS and NMR metabolomics, Metabolomics 10 (2014) 574–588.

[47] A. Scalbert, L. Brennan, O. Fiehn, T. Hankemeier, B. S. Kristal, B. van Ommen, E. P.Guillot, E. Verheij, D. Wishart, S. Wopereis, Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research, Metabolomics 5 (2009) 435–458.

[48] A. Hirayama, M. Wakayama, T. Soga, Metabolome analysis by capillary electrophoresis–mass spectrometry, Trends Anal. Chem 61 (2014) 215–222.

[49] H. Mischak1, J. J. Coon, J. Novak, E. M. Weissinger, J. Schanstra, A. F. Dominiczak, Capillary electrophoresis–mass spectrometry as a powerful tool in biomarker discovery and clinical diagnosis: an update of recent developments, Mass Spectrom. Rev. 28 (2009) 703–724.

[50] S. G. Villas-Bôas, D. G. Delicado, M. Åkesson, J. Nielsen, Simultaneous analysis of amino and nonamino organic acids as methyl chloroformate derivatives using gas chromatography-mass spectrometry, Anal. Biochem. 322 (2003) 134-138.

[51] C. Cipollinaten, A. Pierick, A. B. Canelas, R. M. Seifar, A. J. A. van Maris, J. C. van Dam, J. J. Heijnen, A comprehensive method for the quantification of the non-oxidative pentose phosphate pathway intermediates in *Saccharomyces cerevisiae* by GC-IDMS, J. Chromatogr. B 877 (2009) 3231-3236.

[52] C. A. Sellick, R. Hansen, G. M Stephens, R. Goodacre, A. J Dickson, Metabolite extraction from suspension-cultured mammalian cells for global metabolite profiling Nat. Protoc. 6 (2011) 1241-1249.

[53] T. Soga , Y. Ueno, H. Naraoka, Y. Ohashi, M. Tomita, T. Nishioka, Simultaneous determination of anionic intermediates for *Bacillus subtilis* metabolic pathways by capillary electrophoresis electrospray ionization mass spectrometry, Anal. Chem. 74 (2002) 2233–2239.

[54] T. Soga, Y. Ohashi, Y. Ueno, H. Naraoka, M. Tomita, T. Nishioka, Quantitative metabolome analysis using capillary electrophoresis mass spectrometry, J. Proteome. Res. 2 (2003) 488-494.

[55] W. Lu, E.Kimball, J. D. Rabinowitz, A High-Performance Liquid Chromatography-Tandem Mass Spectrometry Method for Quantitation of Nitrogen-Containing Intracellular Metabolites, J. Am. Soc. Mass Spectrom. 17 (2006) 37–50.

[56] J. M. Knee, T. Z. Rzezniczak, A. Barsch, K. Z. Guo, T. J.S. Merritt, A novel ion pairing LC/MS metabolomics protocol for study of a variety of biologically relevant polar metabolites, J. Chromatogr. B 936 (2013) 63– 73.

[57] N. B. Bennette, J. F. Eng, G. C. Dismukes, An LC-MS-based chemical and analytical method for targeted metabolite quantification in the model cyanobacterium Synechococcus sp. PCC 7002, Anal. Chem. 83 (2011) 3808–3816.

[58] A. Periat, I. Kohler, A. Bugey, S. Bieri, F. Versace, C. Stauba, D. Guillarme, Hydrophilic interaction chromatography versus reversed phase liquid chromatography coupled to mass spectrometry: Effect of electrospray ionization source geometry on sensitivity, J. Chromatogr. A 1356 (2014) 211–220.

[59] S. U. Bajad, W. Lu, E. H. Kimball, J. Yuan, C. Peterson, J. D. Rabinowitz, Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry, J. Chromatogr. A 1125 (2006) 76–88.

[60] L. R. Snyder, J. J. Kirkland, J. W. Dolan, Introduction to modern liquid chromatography, John Wiley & Sons, New Jersey, 2010, p. 329-349.

[61] R. M. Seifar, C. Ras, J. C. van Dama, W. M. van Gulik, J. J. Heijnen, W. A. van Winden, Simultaneous quantification of free nucleotides in complex biological samples using ion pair reversed phase liquid chromatography isotope dilution tandem mass spectrometry, Anal. Biochem. 388 (2009) 213–219.

[62] J. M. Buescher, S. Moco, U. Sauer, N. Zamboni, Ultrahigh performance liquid chromatography-tandem mass spectrometry method for fast and robust quantification of anionic and aromatic metabolites, Anal. Chem. 82 (2010) 4403–4412.

[63] L. Coulier, R. Bas, S. Jespersen, E. Verheij, M. J. van der Werf, T. Hankemeier, Simultaneous quantitative analysis of metabolites using ion-pair liquid chromatography-electrospray ionization mass spectrometry, Anal. Chem. 78 (2006) 6573-6582.

[64] W. B. Dunn, A. Erban, R.J. M. Weber, D. J. Creek, M. Brown, R. Breitling, T. Hankemeier, R. Goodacre, S. Neumann, J. Kopka, M. R. Viant, Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics, Metabolomics 9 (2013), S44–S66.

[65] J. J. Pitt, Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry, Clin. Biochem. Rev. 30 (2009) 19-34.

[66] B. K. Matuszewski, M. L. Constanzer, and C. M. Chavez-Eng, Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS, Anal. Chem. 75 (2003) 3019-3030.

[67] E. Stokvis, H. Rosing, J. H. Beijnen, Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? Rapid Commun. Mass Spectrom. 19 (2005) 401–407.

[68] S. Bajad, V. Shulaev, Highly-parallel metabolomics approaches using LC-MS2 for pharmaceutical and environmental analysis, Trends Anal. Chem. 26 (2007) 625–636.

[69] Agilent Technologies, Inc. 2010, Agilent 6400 Series Triple Quad LC/MS System, Concepts Guide, Manual Part Number G3335-90091.

[70] Agilent Technologies, Inc. 2012, 6200Series TOF and 6500 Series Q-TOF LC/MS System, Concepts Guide, Manual Part Number G3335-90142.

[71] X. Ding, H. Ghobarah, X. Zhang, A. Jaochico, X. Liu, G. Deshmukh, B. M. Liederer, C. E. C. A. Hop, B. Dean, High-throughput liquid chromatography/mass spectrometry method for the quantitation of small molecules using accurate mass technologies in supporting discovery drug screening, Rapid Commun. Mass Spectrom. 27 (2013) 401–408.

[72] L. Vaclavik, A. J. Krynitsky, J. I. Rader, Targeted analysis of multiple pharmaceuticals, plant toxins and other secondary metabolites in herbal dietary supplements by ultra-high performance liquid chromatography–quadrupole-orbital ion trap mass spectrometry, Anal. Chim. Acta 810 (2014) 45–60.

[73] F. Schädel, E. F. Lara, Rapid sampling devices for metabolic engineering applications, Appl. Microbiol. Biotechnol. 83 (2009) 199-208.

[74] F. H. Sklad, K. L. McKee, Adenylate Energy Charge (AEC) Response to stress and extraction technique in the Louisiana crayfish, *Procambarus clarkii*, Bull. Environ. Contam. Toxicol. 33 (1984) 584-591.

[75] C. J. Bolten, P. Kiefer, F. Letisse, J. C. Portais, C. Wittmann, Sampling for Metabolome Analysis of Microorganisms, Anal. Chem. 79 (2007) 3843-3849.

[76] O. Magdenoska, J. Martinussen, J. Thykaer, K. Fog Nielsen, Dispersive solid phase extraction combined with ion-pair ultra high-performance liquid chromatography tandem mass spectrometry for quantification of nucleotides in *Lactococcus lactis*, Anal. Biochem. 440 (2013) 166–177

[77] C. Cordeiro, A.P. Freire, Methylglyoxal assay in cells as 2-methylquinoxaline using 1,2-diaminobenzene as derivatizing reagent, Anal. Biochem. 234 (1996) 221–224.

[78] G. Larsson, M. Törnkvist, Rapid sampling, cell inactivation and evaluation of low extracellular glucose concentrations during fed-batch cultivation, J. Biotechnol. 20 (1996) 69-82.

[79] S. G. Villas-Boas, J. H. Pedersen, M. Akesson, J. Smedsgaard, J. Nielsen, Global metabolite analysis of yeast: evaluation of sample preparation methods Yeast, 22 (2005) 1155-1169.

[80] R. M. Seifar, C. Ras, A. T. Deshmukh, K. M. Bekers, C. A. Suarez-Mendez, A. L.B. da Cruz, W. M. van Gulik, J. J. Heijnen, Quantitative analysis of intracellular

coenzymes in *Saccharomyces cerevisiae* using ion pair reversed phase ultra high performance liquid chromatography tandem mass spectrometry, J. Chromatogr. A 1311 (2013) 115– 120.

[81] A. B. Canelas, C. Ras, A. ten Pierick, J. C. van Dam, J. J. Heijnen, W. M. van Gulik, Leakage-free rapid quenching technique for yeast metabolomics, Metabolomics 4 (2008) 226–239.

[82] H. T. Nikerel, M. de Mey, C. Ras, A. ten Pierick , R. M. Seifar, J. C. van Dam, J. J. Heijnen, W. M. van Gulik, Development and application of a differential method for reliable metabolome analysis in *Escherichia coli*, Anal. Biochem. 386 (2009) 9–19.

[83] U. Nasution, W.M. van Gulik, R.J. Kleijn, W.A. van Winden, A. Proell, J.J. Heijnen, Measurement of intracellular metabolites of primary metabolism and adenine nucleotides in chemostat cultivated *Penicillium chrysogenum*, Biotechnol. Bioeng, 94 (2006) 159-166.

[84] S. G. Villas-Bóas, P. Bruheim, Cold glycerol–saline: The promising quenching solution for accurate intracellular metabolite analysis of microbial cells, Anal. Biochem. 370 (2007) 87–97.

[85] H. Link, B. Anselment, D. W. Botz, Leakage of adenylates during cold methanol/glycerol quenching of Escherichia coli, Metabolomics 4 (2008) 240–247.

[86] J. Wahrheit, J. Niklas, E. Heinzle, Evaluation of sampling and quenching procedures for the analysis of intracellular metabolites in CHO suspension cells BMC Proc. 5 (2011) 82-83.

[87] K. Ortmayr, J. Nocon, B. Gasser, D. Mattanovich, S. Hann, G. Koellensperger, Sample preparation workflow for the liquid chromatography tandem mass spectrometry based analysis of nicotinamide adenine dinucleotide phosphate cofactors in yeast, J. Sep. Sci. 37 (2014) 2185-2191.

[88] J. L. Sporty, M. M.Kabir, K. W. Turteltaub, T. Ognibene, S.-J. Lin, G. Bench, Single sample extraction protocol for the quantification of NAD and NADH redox states in *Saccharomyces cerevisiae*, J. Sep. Sci. 31 (2008) 3202-3211.

[89] M. J. Sáez, R. Lagunas, Determination of intermediary metabolites in yeast. Critical examination of the effect of sampling conditions and recommendations for obtaining true levels. Mol. Cell Biochem. 13 (1976) 73–78.

[90] W. de Koning, K. van Dam, A method for the determinations of changes of glycolytic metabolites in yeast on a sub second time scale using extraction at neutral pH. Anal. Biochem. 204 (1992) 118–123.

[91] C. Wittmann, J. O. Krömer, P. Kiefer, T. Binz, E. Heinzle, Impact of the cold shock phenomenon on quantifcation of intracellular metabolites in bacteria, Anal. Biochem. 327 (2004) 135–139.

[92] C. B. Jendresen, M. Kilstrup, J. Martinussen, A simplified method for rapid quantification of intracellular nucleoside triphosphates by one-dimensional thin-layer chromatography, Anal. Biochem. 409 (2011) 249–259.

[93] M. Volmer, S. Northoff, S. Scholz, T. Thüte, H. Büntemeyer, T. Noll, Fast filtration for metabolome sampling of suspended animal cells, Biotechnol. Lett. 33 (2011) 495–502

[94] Fiehn O., Metabolomics—The link between genotypes and phenotypes. Plant Mol. Biol. 48 (2002) 155–171.

[95] L.P. de Jonge, R. D. Douma, J. J. Heijnen, W. M. van Gulik, Optimization of cold methanol quenching for quantitative metabolomics of *Penicillium chrysogenum*, Metabolomics 8 (2012) 727–735.

[96] J. Spura, L. C. Reimer, P. Wieloch, K. Schreiber, S. Buchinger, D. Schomburg, A method for enzyme quenching in microbial metabolome analysis successfully applied to gram-positive and gram-negative bacteria and yeast, Anal. Biochem. 394 (2009) 192-201.

[97] Y. Kassama, Y Xu, W B. Dunn, N Geukens, J Annė, R. Goodacre, Assessment of adaptive focused acoustics versus manual vortex/freeze-thaw for intracellular metabolite extraction from *Streptomyces lividans* producing recombinant proteins using GC-MS and multi-block principal component analysis, Analyst 135 (2010) 934–942.

[98] A. Buchholz, R. Takors, C. Wandrey, Quantification of intracellular metabolites in *Escherichia Coli* K12 using liquid chromatographic-electrospray ionization tandem mass spectrometric techniques, Anal. Biochem. 295 (2001) 129–137.

[99] A. Buchholz, J. Hurlebaus, C. Wandrey, R. Takors, Metabolomics: quantification of intracellular metabolite dynamics, Biomol. Eng. 19 (2002) 5–15.

[100] M. H. Buckstein, J. He, H. Rubin, Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*, J. Bacteriol. 190 (2008) 718–726

[101] A. B. Canelas, A. ten Pierick, C. Ras, R. M. Seifar, J. C. van Dam, W. M. van Gulik, J. J. Heijnen, Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics, Anal. Chem. 81 (2009) 7379–7389.

[102] M. Faijes**,** A. E Mars**,** E. J Smid, Comparison of quenching and extraction methodologies for metabolome analysis of *Lactobacillus plantarum, Microb. Cell Fact.* **6 (**2007) 1-8.

[103] J. Czarnecka, M. Cieślakb, K. Michał, Application of solid phase extraction and high-performance liquid chromatography to qualitative and quantitative analysis of nucleotides and nucleosides in human cerebrospinal fluid, J. Chromatogr. B 822 (2005) 85–90.

[104] S. Yang, R. E. Synovec, Marina G. Kalyuzhnaya, M. E. Lidstrom, Development of a solid phase extraction protocol coupled with liquid chromatography mass spectrometry to analyze central carbon metabolites in lake sediment microcosms, J. Sep. Sci. 34 (2011) 3597–3605.

[105] S. Deshmukha, A. Frolov, A. Marcillo, C. Birkemeyer, Selective removal of phosphate for analysis of organic acids in complex samples, J. Chromatogr. A 1388 (2015) 1–8.

[106] V. Bezy1, P. Chaimbault, P. Morin, S. E. Unger, M. C. Bernard, L. A. Agrofoglio, Analysis and validation of the phosphorylated metabolites of two anti-human immunodeficiency virus nucleotides (stavudine and didanosine) by pressure-assisted CE-ESI-MS/MS in cell extracts: Sensitivity enhancement by the use of perfluorinated acids and alcohols as coaxial sheath-liquidmake-up constituents, Electrophoresis 27 (2006) 2464–2476.

[107] N. C. van de Merbel, Quantitative determination of endogenous compounds in biological samples using chromatographic techniques, Trends Anal. Chem. 27 (2008) 924-933.

[108] D. Tsikas, A proposal for comparing methods of quantitative analysis of endogenous compounds in biological systems by using the relative lower limit of quantification (rLLOQ), J. Chromatogr. B 877 (2009) 2244–2251.

[109] S. Naz, M. Vallejo, A. García, C. Barbas, Method validation strategies involved in non-targeted metabolomics, J. Chromatogr. A 1353 (2014) 99–105.

[110] R. K. Boyd, C. Basic, R. A. Bethem, Trace Quantitative Analysis by Mass Spectrometry, John Wiley & Sons, Chichester, England, 2008, p. 44-46

[111] R. K. Boyd, C. Basic, R. A. Bethem, Trace Quantitative Analysis by Mass Spectrometry, John Wiley & Sons, Chichester, England, 2008, p. 419-427.

[112] W. Li, L. H. Cohen, Quantitation of endogenous analytes in biofluid without a true blank matrix, Anal. Chem. 75 (2003) 5854-5859.

[113] M. R. Mashego, L. Wu, J. C. Van Dam, C. Ras, J. L. Vinke, W. A. Van Winden, W. M. Van Gulik, J. J. Heijnen, MIRACLE: Mass isotopomer ratio analysis of U-13C-labeled extracts. A new method for accurate quantification of changes in concentrations of intracellular metabolites, Biotechnol. Bioeng. 85 (2004) 621-628.

[114] M. Sugimoto, M. Kawakami, M. Robert, T. Soga1, M. Tomita, Bioinformatics tools for mass spectroscopy-based metabolomic data processing and analysis, Curr. Bioinform. 7 (2012) 96-108.

[115] M. J. Brauer, J. Yuan, B. D. Bennett, W. Lu, E. Kimball, D. Botstein, J. D. Rabinowitz, Conservation of the metabolomic response to starvation across two divergent microbes Proc. Natl. Acad. Sci. 103 (2006) 19302-19307.

[116] Andrea Thorhallsdottir, Master project: Ion chromatography-mass spectrometry for the determination of microbial metabolites, July 2013.

[117] J. A. Ballantine, D. E. Games, P. S. Slater, The enhancement of polyphosphonate salt electrospray mass spectrometry and the high-performance liquid chromatographic separation of triphosphate esters by diethylamine treatment, Rapid Commun. Mass Spectrom. 11 (1997) 624.

[118] M. Wellerdiek, D. Winterhoff, W. Reule, J. Brandner, M. Oldiges, Metabolic quenching of *Corynebacterium glutamicum*: efficiency of methods and impact of cold shock, Bioprocess Biosyst. Eng. 32 (2009) 581–592.

[119] Daniel Killerup Svenssen, Master project: Ion-pair UHPLC-MS/MS and Q-TOF MS analytical methods for intracellular metabolites in yeast and Aspergillus, July 2014.

[120] L. E. Sojo, G. Lum, P. Chee, Internal standard signal suppression by co-eluting analyte in isotope dilution LC-ESI-MS, Analyst 128 (2003) 51–54.

[121] A. Tan, I. A. Lévesque, I. M. Lévesque, F. Viel, N. Boudreau, A. Lévesque, Analyte and internal standard cross signal contributions and their impact on quantitation in LC-MS based bioanalysis, J. Chomatogr. B. 879 (2011) 1954-1960.

[122] J. P. van Dijken, J. Bauer, L. Brambilla, P. Duboc, J. Francois, C. Gancedo, M. Giuseppin, J. Heijnen, M. Hoare, H. Lange, E. Madden, P. Niederberger, J. Nielsen, J. Parrou, T. Petit, D. Porro, M. Reuss, N. van Riel, M. Rizzi, H. Steensma, C. Verrips, J. Vindeløv, J. Pronk, An interlaboratory comparison of physiological and genetic properties of four Saccharomyces cerevisiae strains, Enzyme Microb. Technol. 26 (2000) 706–714.

[123] A. B. Canelas, N. Harrison, A. Fazio, J. Zhang, J. P. Pitkänen, J. van den Brink, B. M. Bakker, L. Bogner, J. Bouwman, J. I. Castrillo, A. Cankorur, P. Chumnanpuen, P. D. Lapujade, D. Dikicioglu, K. van Eunen, J. C. Ewald, J. J. Heijnen, B. Kirdar, I. Mattila, F. I. C. Mensonides, A. Niebel, M. Penttilä, J. T. Pronk, M. Reuss, L. Salusjärvi, U. Sauer, D. Sherman, M. Siemann-Herzberg, H. Westerhoff, J. de Winde, D. Petranovic, S. G. Oliver, C. T. Workman, N. Zamboni, J. Nielsen, Integrated multilaboratory systems biology reveals differences in protein metabolism between two reference yeast strains, Nat. Commun. 1 (2010) 145-153.

[124] J. Martinussen, S. L. L. Wadskov-Hansen, K. Hammer, Two nucleoside uptake systems in *Lactococcus lactis*: competition between purine nucleosides and cytidine allows for modulation of intracellular nucleotide pools, J. Bacteriol. 185 (2003) 1503-1508.

[125] C. M. Waters, W. Lu, J. D. Rabinowitz, B. L. Bassler, Quorum sensing controls biofilm formation in *Vibrio Cholerae* through modulation of cyclic diGMP levels and repression of *vpsT*, J. Bacteriol. 190 (2008) 2527–2536.

[126] A. Klitgaard, A. Iversen, M. R. Andersen, T. O. Larsen, J. C. Frisvad, K. F. Nielsen, Aggressive dereplication using UHPLC-DAD-QTOF: screening extracts for up to 3000 fungal secondary metabolites, Anal. Bioanal. Chem. 406 (2014)1933-1943.

[127] J. Nielsen, S. Oliver, The next wave in metabolome analysis. Trends Biotechnol. 23 (2005) 544–546.

[128] J. L. Griffin, The Cinderella story of metabolic profiling: does metabolomics get to go to the functional genomics ball? Philos. Trans. R. Soc. Lond. B. Biol. Sci. 361 (2006) 147–161.

6. Papers

6.1 Paper 1 - Dispersive solid phase extraction combined with ion-pair ultra high-performance liquid chromatography tandem mass spectrometry for quantification of nucleotides in *Lactococcus lactis*.

Olivera Magdenoska; Jan Martinussen; Jette Thykær; Kristian Fog Nielsen.

Paper published in Analytical Biochemistry 2013

Analytical Biochemistry 440 (2013) 166–177

Contents lists available at SciVerse ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Dispersive solid phase extraction combined with ion-pair ultra high-performance liquid chromatography tandem mass spectrometry for quantification of nucleotides in *Lactococcus lactis*

Olivera Magdenoska [a,*], Jan Martinussen [a], Jette Thykaer [b], Kristian Fog Nielsen [a]

[a] Metabolic Signaling and Regulation Group, Department of Systems Biology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
[b] Fungal Physiology and Biotechnology Group, Department of Systems Biology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 21 December 2012
Received in revised form 23 May 2013
Accepted 24 May 2013
Available online 6 June 2013

Keywords:
Nucleotides
Ion-pair reversed phase liquid chromatography
Isotope dilution tandem mass spectrometry
Dispersive solid phase extraction
Lactococcus lactis

ABSTRACT

Analysis of intracellular metabolites in bacteria is of utmost importance for systems biology and at the same time analytically challenging due to the large difference in concentrations, multiple negative charges, and high polarity of these compounds. To challenge this, a method based on dispersive solid phase extraction with charcoal and subsequent analysis with ion-pair liquid chromatography coupled with electrospray ionization tandem mass spectrometry was established for quantification of intracellular pools of the 28 most important nucleotides. The method can handle extracts where cells leak during the quenching. Using a Phenyl-Hexyl column and tributylamine as volatile ion-pair reagent, sufficient retention and separation was achieved for mono-, di-, and triphosphorylated nucleotides. Stable isotope labeled nucleotides were used as internal standards for some analytes. The method was validated by determination of the recovery, matrix effects, accuracy, linearity, and limit of detection based on spiking of medium blank as well as standard addition to quenched *Lactococcus lactis* samples. For standard addition experiments, the isotope-labeled standards needed to be added in similar or higher concentrations as the analytes. *L. lactis* samples had an energy charge of 0.97 ± 0.001 which was consistent with literature, whereas some differences were observed compared with legacy data based on ^{33}P labeling.

The nucleotides are particularly important parts of the intracellular metabolome not only because of their role as substrates in the synthesis of RNA and DNA but also because they are involved in virtually every metabolic pathway either directly as providers of energy or as allosteric effectors. Cyclic nucleotides have been shown to be a part of many signaling pathways, guanine nucleotides are involved in protein synthesis, and adenosine 5′-triphosphate serves as a primary energy supply for transport, cell motion, and many biosynthetic processes [1–4]. Thus, knowing the intracellular concentration of nucleotides is important for understanding many biological processes within the cell.

For quantitative measurement of nucleotide pools, the sample preparation is of utmost importance because instant quenching of the cell metabolism is required due to the very fast turnover rates of the nucleotide pools within the cell [5,6]. This step can easily be a source of errors due to the possible enzymatic change of nucleotide pools as well as cell leakage in cases where the cells are separated from the growth medium [7]. Thus, merging the quenching and extraction steps is a solution for leaky cells, as shown by Martinussen and coworkers [8], where formic acid was added to *Lactococcus lactis* cells and subsequently three freeze–thaw cycles were used to extract the intracellular metabolites [9].

Purification of the extracted intracellular metabolites is a difficult task, considering both the wide range of concentrations and the diverse physiochemical properties of the intracellular metabolites. Solid phase extraction (SPE)[1] is a widely used method for either sample cleanup or trace enrichment in a variety of matrices, from environmental to biological samples [10–12]. Anastassiades and coworkers [13] proposed an SPE method for cleanup of food and environmental samples called dispersive solid phase extraction (DSPE). In DSPE, the sorbent material is added to the extract to separate the compounds of interest from the matrix components and then is removed from the extract by centrifugation. One of the main advantages of using DSPE is the small amounts of sorbent and solvent that are required, reducing handling and costs [11]. In this regard, charcoal has been proven to be a suitable sorbent that is able

* Corresponding author. Fax: +45 45 88 41 48.
E-mail address: olima@bio.dtu.dk (O. Magdenoska).

[1] *Abbreviations used:* SPE, solid phase extraction; DSPE, dispersive solid phase extraction; TLC, thin-layer chromatography; LC, liquid chromatography; MS, mass spectrometry; MS/MS, tandem mass spectrometry; TBA, tributylamine; HCl, hydrochloric acid; DBAA, dibutylamine acetate; ESI, electrospray ionization; IS, internal standard; SA, synthetic amino acid; UHPLC, ultra high-performance liquid chromatography; QTOF, quadrupole time-of-flight; LOD, limit of detection; RSD, relative standard deviation.

0003-2697/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ab.2013.05.023

to retain nucleotides from biological extracts [9,14]. Nucleotide adsorption on charcoal is based on interaction of the aromatic ring electrons from the purine or pyrimidine moiety with the π electrons from the charcoal.

Quantification methods for nucleotides, given in the literature, involve incorporation of radioactive phosphate followed by one- or two-dimensional thin-layer chromatography (TLC) [9,15] as well as less sensitive and selective ion-pair, ion-exchange, and ion chromatography with fluorescence, ultraviolet, and conductivity detection [16–19]. In addition, the combination of liquid chromatography (LC) and mass spectrometry (MS) has been shown to be a powerful analytical technique for quantitative analysis of nucleotides in *Saccharomyces cerevisiae*, *Bacillus subtilis*, *Escherichia coli*, human plasma, and animal tissue [20–23]. In addition, LC–MS/MS (tandem mass spectrometry) with porous graphitic carbon as a stationary phase has been used for analysis of araCTP, CTP, and dCTP [24]. LC–MS/MS fulfills three important requirements when dealing with analysis of intracellular metabolites: (i) sensitivity, (ii) wide linear range, and (iii) no need to form thermally stable derivatives of the tri-, di-, or monophosphorylated nucleotides as required for gas chromatography mass spectrometry [6,25]. Due to the high polarity and multiple charges, nucleotides are not well retained under reversed phase conditions, and therefore methods such as hydrophilic interaction chromatography (HILIC) and ion-exchange chromatography are needed for their separation [26,27]. A subtype of the latter is ion-pair chromatography that is made by dynamic modification of a reversed phase separation by coating the surface with charged but still hydrophobic ion-pair reagent added to the mobile phase [20,21,28,29]. Ion-pair chromatography is a promising alternative especially when dealing with isomeric compounds that any other chromatography technique fails to separate. Furthermore, the hydrophobic parts of the nucleotides can also interact with the hydrophobic parts of the reversed phase column, thereby giving mixed-mode conditions. For compatibility with atmospheric ionization techniques used for LC–MS, volatile ion-pairing reagents such as tributyl-, dibutyl-, and triethylamine are required because the ion source will clog within minutes of operation if a nonvolatile reagent such as tetrabutylamine is used [30].

When using LC–MS as a detection technique for analysis of complex biological extracts, enhancement or suppression of the ionization is often observed, resulting in different responses for the compounds of interest than are seen from pure standard solutions [31,32]. This occurs when target compounds coelute with matrix interferences and, therefore, the modulation of the LC–MS signal (i) affects reproducibility, (ii) leads to systematic errors, and (iii) can obscure the detection of the target compound in extreme cases.

To prevent the signal modulation especially observed in electrospray ionization (ESI), four alternatives exist: (i) development of more selective chromatographic separation, (ii) better sample purification (e.g., by SPE), (iii) use of more sensitive MS instruments that allow sample dilution, and (iv) use of stable isotope-labeled internal standards (ISs) that can compensate for matrix effects [33]. Due to the similar physicochemical properties, the isotope-labeled ISs also correct for losses and decomposition during sample preparation.

In this study, we present an ion-pair LC isotope dilution MS/MS method combined with charcoal sample purifications for analysis of nucleotide pools. Tributylamine (TBA) was used as ion-pair reagent to modify the mobile phase and facilitate the retention of the nucleotides. Subsequently the applicability of the method was demonstrated by analysis of nucleotide pools in *L. lactis*. To the best of the authors' knowledge, there is no previous report regarding the combination of DSPE purification using charcoal and ion-pair LC–MS/MS analysis for determination of nucleotide pool sizes.

Materials and methods

Materials

Nucleotide standards were purchased from Sigma–Aldrich (Steinheim, Germany). Uniformly isotope-labeled (^{13}C and/or ^{15}N) nucleotides were used as ISs and were purchased from Silantes (Munich, Germany) and Sigma–Aldrich. The chemical purity and isotope enrichment of the isotope-labeled nucleotides were more than 90 and 98%, respectively, except for [U-^{13}C]ATP, [U-^{13}C]GTP, and [U-$^{13}C^{15}N$]CTP, which had a chemical purity of 95%. Amino acids, vitamins, glucose, inorganic salts, active charcoal (C3345), hydrochloric acid (HCl), TBA (puriss. plus grade), 0.5 M dibutylamine acetate (DBAA) concentrate (LC–MS grade), ethanol, methanol, acetonitrile, and acetic acid (LC–MS grade) were obtained from Sigma–Aldrich. Water was purified using a Milli-Q system (Millipore, Bedford, MA, USA).

Stock solutions

Stock solutions of the nucleotides with concentration of 1 mg/ml were prepared in water and stored under –20 °C until use. Aliquots of the stock solutions were used to prepare the daily working solutions by further dilution in 10 mM TBA and 10 mM acetic acid solution. Stock solutions (1 and/or 0.2 mg/ml) of labeled ISs were prepared in water and kept under –20 °C until use. The concentration of the labeled nucleotides used to prepare the IS mixture for spiking the calibration standards and the extracts is given in Table 1.

Cell growth and sampling

L. lactis wild-type strain was grown overnight at 30 °C on agar plates containing M17 broth medium obtained from Oxoid supplied with 1% glucose (GM17) [34]. Ten single colonies were used to inoculate synthetic amino acid (SA) medium [35] containing 0.5% glucose. The culture was subjected to 10-fold dilution series (up to 10^{-7}) and grown overnight at 30 °C. Then 50 ml of culture from the dilution where the cells were still in exponential growth was transferred into 500 ml of preheated SA medium at 30 °C. The OD_{450} of the cultures was measured continuously, starting from 0.03 until 0.5, when 5 ml of culture was quenched with 1 ml of 10 M cold formic acid, followed by the addition of 65 μl of IS mixture. After vigorous mixing, the cultures were stored at –80 °C. For extraction of the metabolites, the cultures were subjected to three freeze–thaw cycles by placing the samples from a –80 °C freezer to an icewater bath, followed by mixing. Then the samples were centrifuged at 4248*g* and the supernatant (6.0 ml) was transferred to a chilled Falcon tube.

Charcoal sample cleanup

Activated charcoal (0.75 g) was suspended in 5 ml of 96% ethanol and 45 ml of water and was vortexed. The suspension was centrifuged for 20 min at 4248*g* and 4 °C, and the supernatant was discarded. The pellet was then mixed with 1.5 ml of 1 M HCl and centrifuged. After discarding the supernatant, the charcoal pellet was washed with 10 ml of water three times and finally resuspended in 1 ml of water.

The quenched bacterial sample was transferred to a chilled tube containing 0.6 ml of the activated charcoal suspension and was vigorously mixed and kept on ice. The suspension was centrifuged for 20 min at 4248*g* and 4 °C. The supernatant was transferred to a chilled Falcon tube and stored at –80 °C. The charcoal pellet was washed two times with ice-cold water and centrifuged for

Table 1
Precursor and product ions of the labeled and unlabeled standards used in the analysis.

Nucleotide	RT(min)	Precursor ion (*m*/*z*)	Product ion (*m*/*z*)	Collision energy (V)	Fragmentor voltage (V)	Cell acceleration voltage (V)	Labeled nucleotides	Precursor ion (*m*/*z*)	Product ion (*m*/*z*)	Concentration (μg/ml)[a]
AMP	22.1	346	79	20	130	3	[U-^{15}N]AMP	350.8	79	2.5
ATP	26	505.9	273	30	95	3	[U-^{13}C]ATP	515.9	278	100
CMP	18.5	322	79	40	120	4	[U-^{13}C^{15}N]CMP	334	79	2.5
CTP	25.6	481.9	159	30	125	4	[U-^{13}C^{15}N]CTP	493.9	159	6
dATP	26.2	489.9	158.9	35	130	3	[U-^{15}N]dATP	495	158.7	2
dCMP	19.8	306	79	40	130	4	[U-^{13}C^{15}N]dCMP	318	79	1.25
dCTP	25.6	465.9	158.9	25	120	3	[U-^{13}C^{15}N]dCTP	477.9	158.9	5
dGMP	21.8	346	79	30	90	4	[U-^{13}C^{15}N]dGMP	361.1	79	2
dGTP	25.8	505.9	257	35	115	3	[U-^{13}C^{15}N]dGTP	521.1	262	2.5
TMP	21.9	321	124.9	20	110	4	[U-^{13}C^{15}N]TMP	333	131.8	1.5
GMP	20.8	362	79	25	110	4	[U-^{13}C^{15}N]GMP	377.1	79	2.5
GTP	25.7	521.9	159	35	100	3	[U-^{13}C]GTP	531.9	159	60
TTP	26	480.9	159	25	100	3	[U-^{13}C^{15}N]TTP	493	159	1.5
UMP	20	323	79	40	100	4	[U-^{13}C^{15}N]UMP	334	79	2.5
UTP	25.8	482.9	159	30	95	4	[U-^{15}N]UTP	485.1	158.9	50
ADP	24.7	425.9	134	25	115	3	–	–	–	–
cAMP	21.1	328.1	134	25	115	4	–	–	–	–
c-diGMP	23.5	689	149.9	35	100	4	–	–	–	–
CDP	24.4	401.9	79	40	90	4	–	–	–	–
cGMP	19.3	344	150	25	115	3	–	–	–	–
dUMP	21.1	307	195	15	110	4	–	–	–	–
IMP	21	347	79	35	115	4	–	–	–	–
OMP	24.4	367	79	25	95	3	–	–	–	–
UDP-Glc	21.8	564.9	323	30	120	3	–	–	–	–
UDP	24.4	402.9	79	40	115	3	–	–	–	–
XMP	24.5	363	211.1	20	100	4	–	–	–	–
ZMP	20.7	337	78.9	35	115	3	–	–	–	–
GDP	24.3	442	159	30	120	3	–	–	–	–

Note: Retention time (RT), optimized collision energy, fragmentor voltage, and cell acceleration voltage of the multiple reaction monitoring transitions are also given in the table. c-di-GMP, cyclic diguanylate; UDP-Glc, uridine diphosphate glucose; GTP, guanosine-5′-triphosphate; ATP, adenosine-5′-triphosphate; dGTP, deoxyguanosine triphosphate; dATP, deoxyadenosine triphosphate; UTP, uridine-5′-triphosphate; CTP, cytidine-5′-triphosphate; TTP, thymidine triphosphate; dCTP, deoxycytidine triphosphate; GDP, guanosine-5′-diphosphate; ADP, adenosine-5′-diphosphate; UDP, uridine-5′-diphosphate; CDP, cytidine-5′-diphosphate; OMP, orotidine-5′-monophosphate; XMP, xanthosine monophosphate; GMP, guanosine-5′-monophosphate; IMP, inosine-5′-monophosphate; AMP, adenosine-5′-monophosphate; dGMP, deoxyguanosine monophosphate; cGMP, guanosine cyclic-3′,5′-monophosphate; ZMP, 5-aminoimidazole-4-carboxamide ribotide; cAMP, adenosine cyclic-3′,5′-monophosphate; UMP, uridine-5′-monophosphate; CMP, cytidine-5′-monophosphate; TMP, thymidine monophosphate; dUMP, deoxyuridine monophosphate; dCMP, deoxycytidine monophosphate. The experimental details are given in Materials and Methods.

[a] Concentration of the labeled nucleotides in the IS mixture used to spike the calibration standards and the extracts for the quantification experiments.

20 min at 4248*g* and 4 °C. The nucleotides were eluted from the charcoal with ice-cold solvent containing 2% NH_3 and 50% acetonitrile in water. The samples were evaporated under nitrogen and finally resuspended in 325 μl of 10 mM TBA and 10 mM acetic acid solution.

UHPLC–MS/MS analysis

Analysis was performed on an Agilent 1290 binary ultra high-performance liquid chromatography (UHPLC) system coupled with an Agilent 6460 triple quadrupole system (Torrance, CA, USA) equipped with an Agilent jet stream ESI source and was operated in negative ion mode. Nitrogen was used as collision gas. The source and fragmentation parameter were optimized for each of the nucleotides using 10 μg/ml single standard dissolved in mobile phase A (10 mM TBA containing 10 mM acetic acid), bypassing the column. The most intense product ions were selected by increasing the collision energy voltage in the product ion scan and monitoring the intensity of their MS signal. The optimized ion-source-dependent parameters were as follows: gas temperature, 300 °C; sheath gas temperature, 400 °C; nebulizer gas flow rate, 8 L/min; nebulizer pressure, 50 psi; and capillary voltage, 4500 V. The entrance potential (ΔEMV, electron multiplier voltage) was kept at 500 V for all of the transitions. The mass spectrometer resolution was set to "unit" in both Q_1 (first quadrupole) and Q_3 (third quadrupole). Unless otherwise stated, separation was performed by using an Agilent Poroshell 120 Phenyl-Hexyl column (2.7 μm, 100 × 2.1 mm, operated at 40 °C). An Agilent 1290 Infinity in-line filter (0.3 μm) was used to protect the column. Mobile phase A was 10 mM TBA and 10 mM acetic acid (pH 5.5), and mobile phase B was 90% (v/v) methanol containing 10 mM TBA and 10 mM acetic acid. The gradient used was as follows: 0 to 5 min of 0% B, 5 to 10 min of 0 to 2% B, 10 to 11 min of 2 to 9% B, 11 to 16 min of 9% B, 16 to 24 min of 9 to 50% B, 24 to 28.5 min of 100% B, 28.5 to 30 min of 100% B, 30 to 30.5 min of 100 to 0% B, and 30.5 to 36 min of 0% B. The injection volume was 10 μl. The LC–MS/MS run was divided into three time segments. During the first (0–11 min) and third (32–36 min) time segments, the valve from the mass spectrometer was diverted to waste in order to minimize the contamination of the source. The intracellular concentrations were determined assuming an intracellular volume of 1.67 ml for 1 g (dry weight) of cells [36].

UHPLC accurate mass verification

To investigate sample purity and adduct pattern and to verify the dCTP/CTP ratio, few representative extracts were analyzed by passing the effluent coming from the Agilent 1290 binary UHPLC system to an Agilent 6550 quadrupole time-of-flight (QTOF) instrument operated in ESI^- in the 2-GHz extended dynamic mode at a resolution of 25,000 full width at half-maximum (FWHM). The QTOF instrument was tuned for fragile molecules by lowering all of the potentials in the ion path by 5 V relative to the autotune values.

Method validation and quantification

For calculating the accuracy/recovery and precision, two approaches were used: (i) the standard addition approach for endogenous compounds proposed by Tsikas [37] and (ii) validation based on spiking of the medium used for growing the cells. For the first

Table 2
Accuracy and precision of the standard addition experiment including the determined intracellular nucleotide amounts in *L. lactis*.

	Nucleotide	R^2	LOD (μg)	Measured[a] (μg)	Added amount[b] (μg)					Day 1		Day 2		Day 3		nmol/mg dry weight	Intracellular concentration (μM) [d]
										Accuracy[c](%)	Precision[c](RSD%)	Accuracy (%)	Precision (RSD%)	Accuracy (%)	Precision (RSD%)		
With IS	AMP	0.992	0.004	0.07	**0.05**	**0.11**	**0.16**	**0.22**	**0.27**	87	17	95	9	87	2	0.4	0.2
	ATP	0.995	0.004	4	**1.66**	**3.33**	**4.99**	**6.66**	**8.32**	88	17	88	15	87	15	16	10
	CMP	0.995	0.03	0.01	**0.08**	**0.16**	**0.24**	**0.32**	**0.40**	94	7	84	19	94	6	0.1	0.1
	CTP	0.985	0.001	0.3	**0.11**	**0.21**	**0.32**	**0.43**	**0.53**	88	16	85	17	82	19	1	0.6
	dATP	0.987	0.001	0.1	**0.02**	**0.05**	**0.07**	**0.09**	**0.12**	95	6	95	1	94	10	0.4	0.2
	dCMP	0.994	0.002	0.005	**0.02**	**0.04**	**0.06**	**0.08**	**0.10**	91	13	91	14	88	8	0.03	0.02
	dCTP	0.993	0.003	0.3	**0.17**	**0.33**	**0.50**	**0.67**	**0.83**	93	9	90	11	93	10	1.4	0.9
	dGMP	0.990	0.015	ND	**0.17**	**0.33**	**0.50**	**0.67**	**0.84**	94	4	93	12	94	7	ND	ND
	dGTP	0.95	0.008	0.01	0.10	0.20	**0.29**	**0.39**	**0.49**	92	5	89	17	86	23	0.04	0.03
	dTMP	0.991	0.003	0.05	0.08	**0.15**	**0.23**	**0.30**	**0.38**	95	7	94	8	90	7	0.3	0.2
	GMP	0.992	0.005	0.02	**0.03**	**0.07**	**0.10**	**0.13**	**0.16**	89	2	84	18	89	12	0.1	0.05
	GTP	0.994	0.003	3	**1.68**	**3.36**	**5.04**	**6.72**	**8.40**	93	8	86	11	96	0.4	10	6
	TTP	0.991	0.001	0.2	**0.17**	**0.34**	**0.51**	**0.68**	**0.85**	84	19	92	9	85	11	0.8	0.5
	UMP	0.991	0.001	0.1	**0.09**	**0.17**	**0.26**	**0.35**	**0.43**	86	11	92	14	89	8	0.5	0.3
	UTP	0.992	0.001	1	**0.79**	**1.59**	**2.38**	**3.17**	**3.96**	84	22	98	1	86	13	5	3
Without IS	ADP	0.975	0.004	0.4	0.02	**0.04**	**0.06**	**0.08**	**0.10**	87	2	95	1	99	2	1.7	1
	cAMP	0.987	0.001	ND	0.02	**0.04**	**0.07**	**0.09**	**0.11**	92	14	76	14	89	7	ND	ND
	cdiGMP	0.988	0.001	ND	**0.05**	**0.10**	**0.15**	**0.19**	**0.24**	87	6	91	14	91	11	ND	ND
	CDP	0.99	0.009	0.003	**0.08**	**0.16**	**0.25**	**0.33**	**0.41**	85	19	66	34	84	6	0.02	0.01
	cGMP	0.986	0.001	ND	**0.02**	**0.04**	**0.05**	**0.07**	**0.09**	86	12	85	19	88	14	ND	ND
	dUMP	0.982	0.002	ND	0.08	**0.17**	**0.25**	**0.33**	**0.41**	84	4	84	4	96	8	ND	ND
	IMP	0.798	0.002	0.1	0.05	0.10	**0.15**	**0.20**	**0.25**	79	7	80	2	80	8	0.6	0.4
	OMP	0.981	0.003	ND	**0.06**	**0.12**	**0.17**	**0.23**	**0.29**	85	9	84	15	91	8	ND	ND
	UDP-Glc	Failed	0.00001		0.04	0.08	0.12	0.16	0.20	255	59	115	87	862	173	–	–
	UDP	0.987	0.004	0.04	**0.07**	**0.14**	**0.22**	**0.29**	**0.36**	89	11	95	1	83	21	0.2	0.1
	XMP	0.984	0.002	0.1	0.16	**0.32**	**0.49**	**0.65**	**0.81**	89	3	86	20	86	16	0.5	0.3
	ZMP	0.923	0.09	0.02	0.15	**0.30**	**0.45**	**0.60**	**0.75**	88	16	86	5	85	2	0.12	0.07
	GDP	0.99	0.003	0.002	**0.04**	**0.08**	**0.12**	**0.16**	**0.20**	85	18	91	9	91	13	0.003	0.002

[a] Amount measured in the sample before spiking.
[b] Externally added amounts of a particular nucleotide.
[c] Accuracy and precision of the measurement of the underlined added amount of a particular nucleotide on 3 different days.
[d] Concentration determined assuming an intracellular volume of 1.67 ml for 1 g dry weight of cells [29]. ND not detected.

approach, a mixture containing 28 target nucleotides at different concentrations was prepared in levels of approximately 0, 40, 80, 120, 160, and 200% of the mass (Table 2) determined in *L. lactis* samples via external calibration. For the compounds that were not detected in the *L. lactis* samples, the concentration in the mixture was approximately 50 times the limit of detection (LOD) in spiked growth medium.

For the standard addition approach, a 500-ml batch of *L. lactis* culture was quenched by transferring 20 ml of culture into 100-ml Erlenmeyer flasks containing cold 4 ml of 10 M formic acid, followed by freezing at –80 °C. After thawing, all cultures were pooled together. The pooled culture was divided by pipetting 6 ml into 15-ml Falcon tubes. These cultures were used in the spiking experiments done on 3 different days by spiking 0, 30, 60, 90, 120, and 150 μl (each level in triplicates) of the concentrated nucleotide mixture and 65 μl of IS mixture (Table 1), resulting in 54 validation samples in total.

For the spiked growth medium validation approach, three groups of samples were prepared: (i) in the neat solution (eluent A), (ii) spiking the medium after charcoal purification and (iii) spiking the medium before charcoal purification by adding 65 μl of IS mixture and all target nucleotides to concentrations of 0, 0.1, 0.3, 0.6, 1.2, and 6 μg/ml. The samples were prepared in triplicates and analyzed by UHPLC–MS/MS on 3 different days.

The calibration curves, for the nucleotides for which isotope-labeled analog was added, were constructed by plotting the peak area ratios of the unlabeled nucleotide standard to the corresponding labeled one versus the concentrations of the nucleotide standard added. For the nucleotides where no IS was added, the calibration curves were constructed by plotting the peak area of the nucleotide versus the concentration of the nucleotide standard added. The quantification was done in Excel. The LOD for the nucleotides was determined at a signal-to-noise (S/N) ratio of 5 from samples spiked in the medium, followed by charcoal purification.

Assessment of matrix effects from growth medium

The approach of Matuszewski and coworkers [32] was used for determination of the matrix effects coming from the medium. Two groups of samples were prepared: (i) in the neat solution (eluent A) and (ii) in the medium after charcoal purification (in triplicates) by adding 65 μl of IS mixture and all target nucleotides to concentrations of 0, 0.1, 0.3, 0.6, 1.2, and 6 μg/ml. The samples were analyzed by UHPLC–MS/MS on 3 different days.

The matrix effects were determined by the ratio of the slope of the calibration curve ($1/x$ weighting) of standards in the neat solution and in the medium spiked after charcoal purification.

Assessment of matrix effects of ATP by T-piece infusion test

To study the total ion suppression, a T-piece infusion test was conducted [38] using 10 μg/ml ATP standard prepared in the mobile phase. This was continuously infused in the eluent from the UHPLC using a syringe pump at 5 μl/min. While monitoring the signal from the infusion, three different samples were injected and separated on the final gradient method: (i) mobile phase, (ii) medium used for growing the cells previously purified with charcoal, and (iii) *L. lactis* extract purified with charcoal.

Results and discussion

Optimization of MS parameters

Due to the strongly ionizing TBA in the eluents, it was only possible to operate the mass spectrometer in negative mode, where all target compounds generated $[M–H]^-$ as the most intense ion. The nucleotides with similarities in their structure (e.g., those containing purine base or mono-, di-, or triphosphate groups) showed a common fragmentation pattern. The fragments used for quantification as well as the optimized collision energy, fragmentor, and cell accelerator voltage values for each fragment are given in Table 1.

Optimization of ion-pair chromatography

To find the separation conditions for the nucleotides, two volatile ion-pair reagents were tested: 10 mM DBAA (pH 7.3) and 10 mM TBA containing 10 mM acetic acid (pH 5.5). Both ion-pair reagents gave sufficient retention for both deoxy- and ribonucleotides that increased in the order of monophosphates < diphosphates < triphosphates, as reported previously by others [20]. The number of phosphate groups, interacting with the ion-pair reagent, determined the elution pattern. As expected, DBAA gave less retention of the nucleotides compared with TBA, which was explainable by the one fewer alkyl chain. During method optimization, various reversed columns were also tested: Phenomenex Luna C_{18}(2)-HST and Onyx Monilithic C_{18}, Agilent Zorbax extended C_{18}, and Poroshell 120 Phenyl-Hexyl. It was noted that the Onyx Monolithic column was significantly less retentive than the other columns and gave broader chromatographic peaks. In general, the Poroshell 120 Phenyl-Hexyl column gave better spreading and less tailing of the chromatographic peaks (data not shown) and, thus, was selected as our standard column.

For the separation, Buescher and coworkers' chromatographic gradient [39] was taken as a starting point, but because interference was observed from A+1 isotopomer on compounds with a 1-Da higher mass (e.g., UMP and CMP with $[M–H]^-$ ions of 323 and 322, respectively), a better separation was needed and it was necessary to use a less steep gradient between 16 and 24 min. For total elution of the nucleotides, it was necessary to increase phase B to 100%.

When conducting the experiments with TBA and dibutylamine, it was noticed that the analysis in which TBA was used as ion-pair reagent gave at least a 5-fold increase in intensity of all the nucleotides. To investigate the effect of the pH of the eluents, additional experiments were performed at pHs 7.3 and 5.5 for both dibutylamine and TBA. In these experiments, the concentration of the acid was changed, whereas the concentration of ion-pair reagent was kept constant. As expected by reducing the acetic acid concentration in both TBA and dibutylamine eluent, there was an increase of nucleotide retention due to the reduced competitive effect of the acetate [29]. When compared, the results obtained with dibutylamine at two different pHs showed only minor differences in the nucleotide signal intensities. The decrease in the acetic acid concentration (pH 7.3), when TBA was used as eluent, led to a decrease of the MS signal for the monophosphates and diphosphates when compared with the analysis obtained with pH 5.5 TBA, but better peak shape and increased MS signal for the triphosphates were obtained with pH 7.3 TBA. Due to a lack of reproducibility of the chromatography obtained using pH 7.3 TBA as eluent and the lower signal intensity obtained with dibutylamine as ion-pair reagent, pH 5.5 TBA was chosen as eluent for further experiments.

In addition, the effect of the solvent used to prepare the samples on the UHPLC–MS/MS analysis was investigated. It was noted that when standard solutions were prepared in mobile phase A, a 2-fold increase of the signal intensity for the nucleotides was observed compared with the standards prepared in water/methanol mixture. This effect was more prominent when higher injection volumes were used. Not only an increase of the intensity but also a reduction of the peak tailing was observed when preparing the

standard solutions in mobile phase A, which corresponded to basic chromatographic rules for achieving better peak shape.

Due to their identical elemental composition, ATP and AMP had the same precursor *m*/*z* as dGTP and dGMP, respectively. AMP and dGMP were separated, which was not the case for ATP and dGTP. MS separation was achieved for the latter two by identifying fragments generated as a result of loss of adenine and guanine units, respectively. Fig. 1 shows the chromatograms of the nucleotides obtained by measurement of 1.2 μg/ml standard mixture.

Due to the one mass unit difference between UMP and CMP, it was necessary to achieve baseline separation in order to eliminate the contribution of the $[M–H]^-$ A+1 isotopic ion of CMP (*m*/*z* 322) to the $[M–H]^-$ ion of UMP (*m*/*z* 323). The same was achieved for dTMP (*m*/*z* 321), dUMP (*m*/*z* 307), and dCMP (*m*/*z* 306), IMP (*m*/*z* 347), AMP (*m*/*z* 346), and dGMP (*m*/*z* 346), and XMP (*m*/*z* 363) and GMP (*m*/*z* 362). Sufficient separation was obtained for CTP (*m*/*z* 482) and dTTP (*m*/*z* 481) as well, but not for UTP (*m*/*z* 483) and CTP (*m*/*z* 482). Therefore, inspection of the chromatograms and manual correction of the software automatic integration of the peaks needed to be done. The latter one was applicable for UDP and CDP as well.

Fig. 2 shows the chromatograms of the unlabeled nucleotides from *L. lactis* together with the added isotope-labeled analogs.

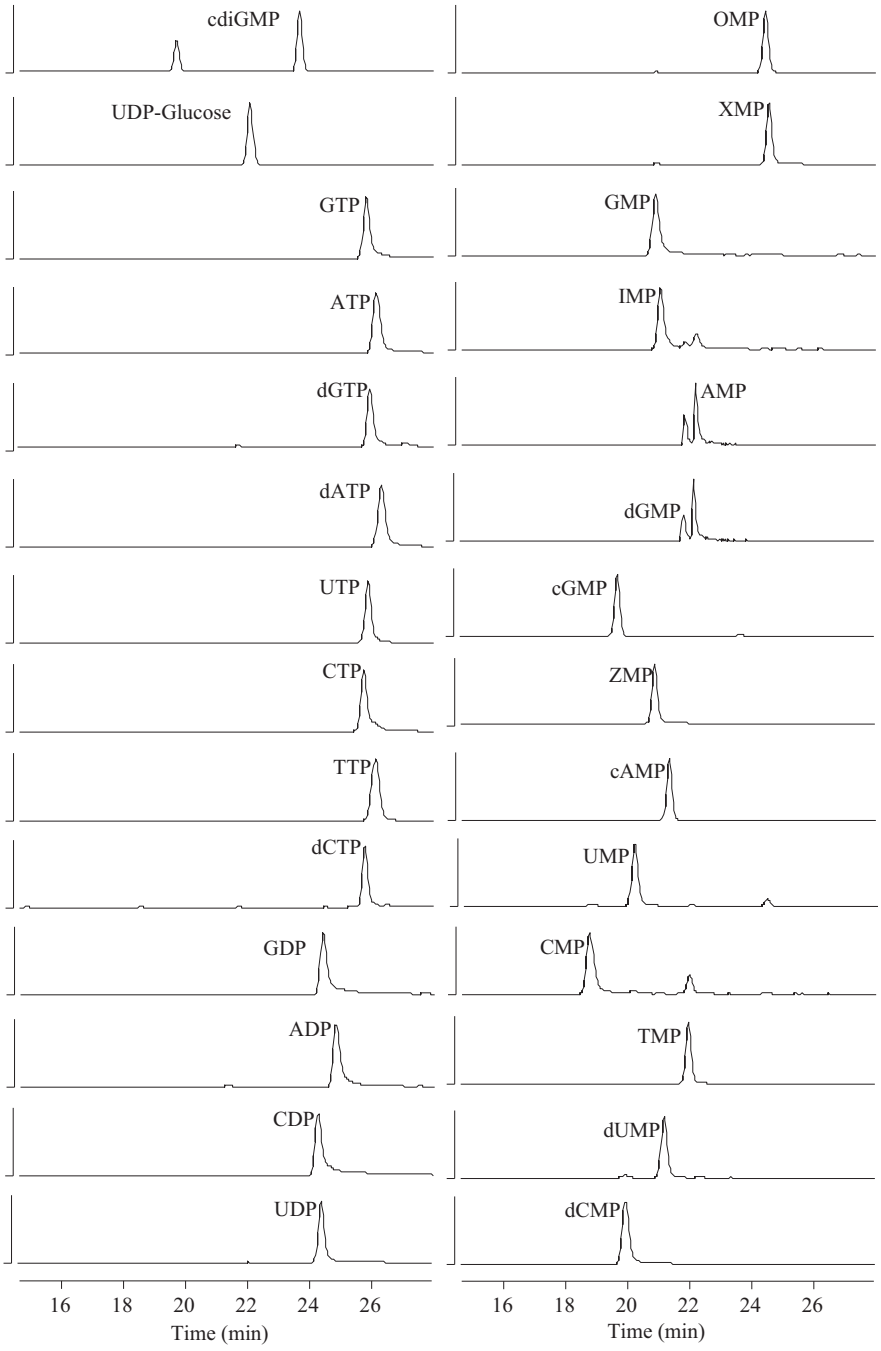


Fig. 1. Chromatograms of the nucleotides obtained from the analysis of 1.2 μg/ml standard mixture. The second time segment, when the nucleotides elute from the column, is shown in the figure. The experimental conditions are given in Materials and Methods. For abbreviations, see note in Table 1.

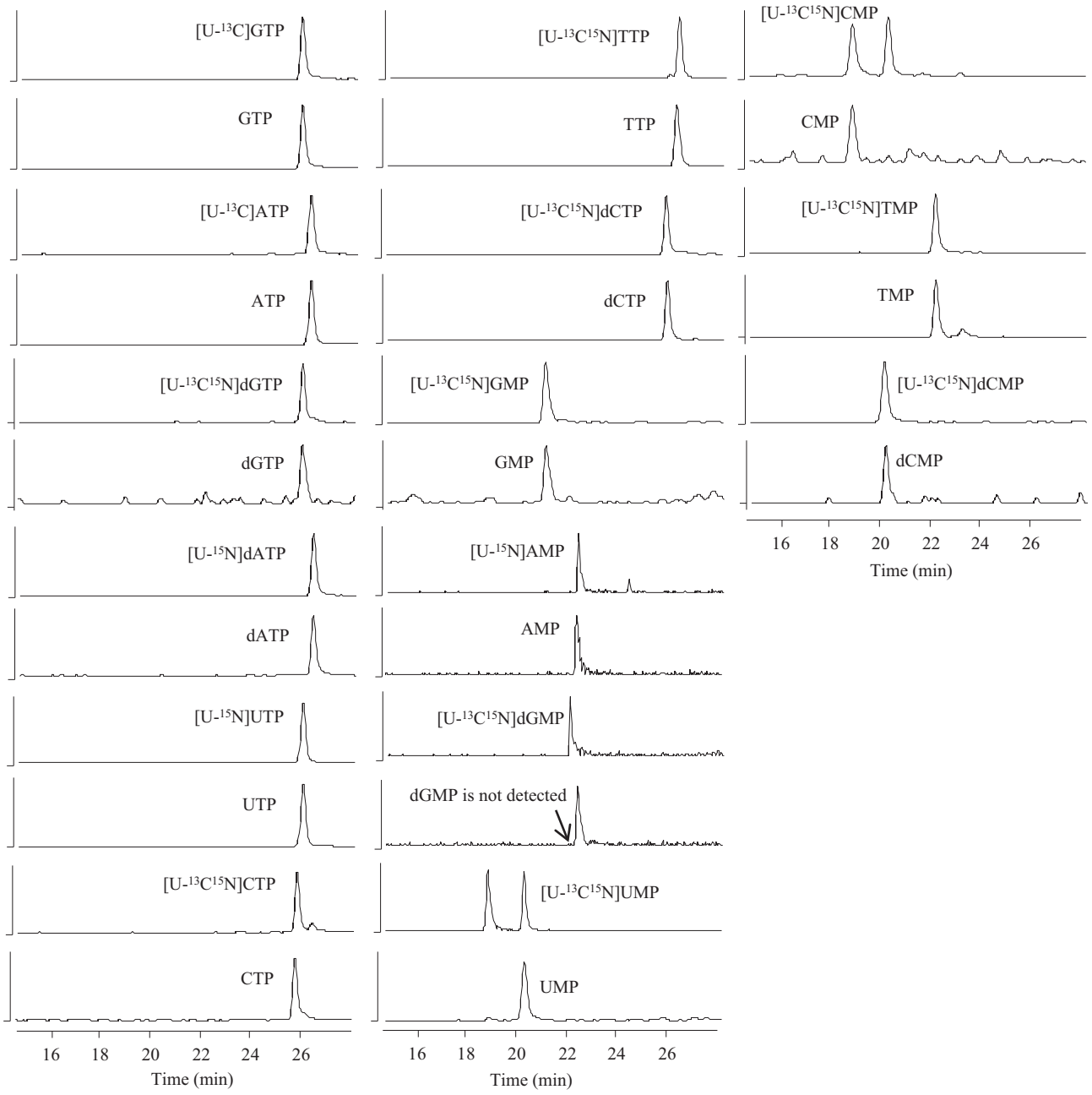


Fig. 2. Chromatograms of the unlabeled nucleotides from *L. lactis* together with the added isotope-labeled analogs.

The retention times of the nucleotides detected in the extract were consistent with those from the standard mixture.

Optimization of sample charcoal purification

Jendresen and coworkers [9] used charcoal for sample purification in combination with TLC detection of radioactively ^{33}P-labeled nucleotides. On the basis of their findings, more detailed investigation was performed on each step of the charcoal cleanup procedure. By analyzing the second step of elution, it was found that only a small amount of the nucleotides remained bound to the charcoal. For ATP, this amount was less than 10% when compared with the whole amount eluted in the first and second elutions (data not shown). To avoid the possible elution of undesired impurities in the second elution step, further experiments were performed with only one elution step. Moreover, to determine the amount of unbound nucleotides, the supernatant was subjected to a second charcoal extraction step. The analysis showed that an insignificant amount (e.g., for ATP, <2% of the total amount detected) of the nucleotides remains in the supernatant unretained by the charcoal (data not shown).

In addition, different organic modifiers such as ethanol, methanol, isopropanol, and acetonitrile in combination with 2% ammonia or 30 mM ion-pair reagent were tested as solvents for eluting the nucleotides from the charcoal. Fig. 3 summarizes the results obtained using different eluents for the elution of the nucleotides from the charcoal. It is shown that in combination with 2% ammonia, 50% (v/v) acetonitrile showed slightly better elution power when compared with the other organic modifiers. For example, 15% more ATP was eluted from the charcoal by using 2% ammonia and 50% acetonitrile compared with the ethanol/ammonia mixture. On the other hand, the combination of organic modifier and ion-

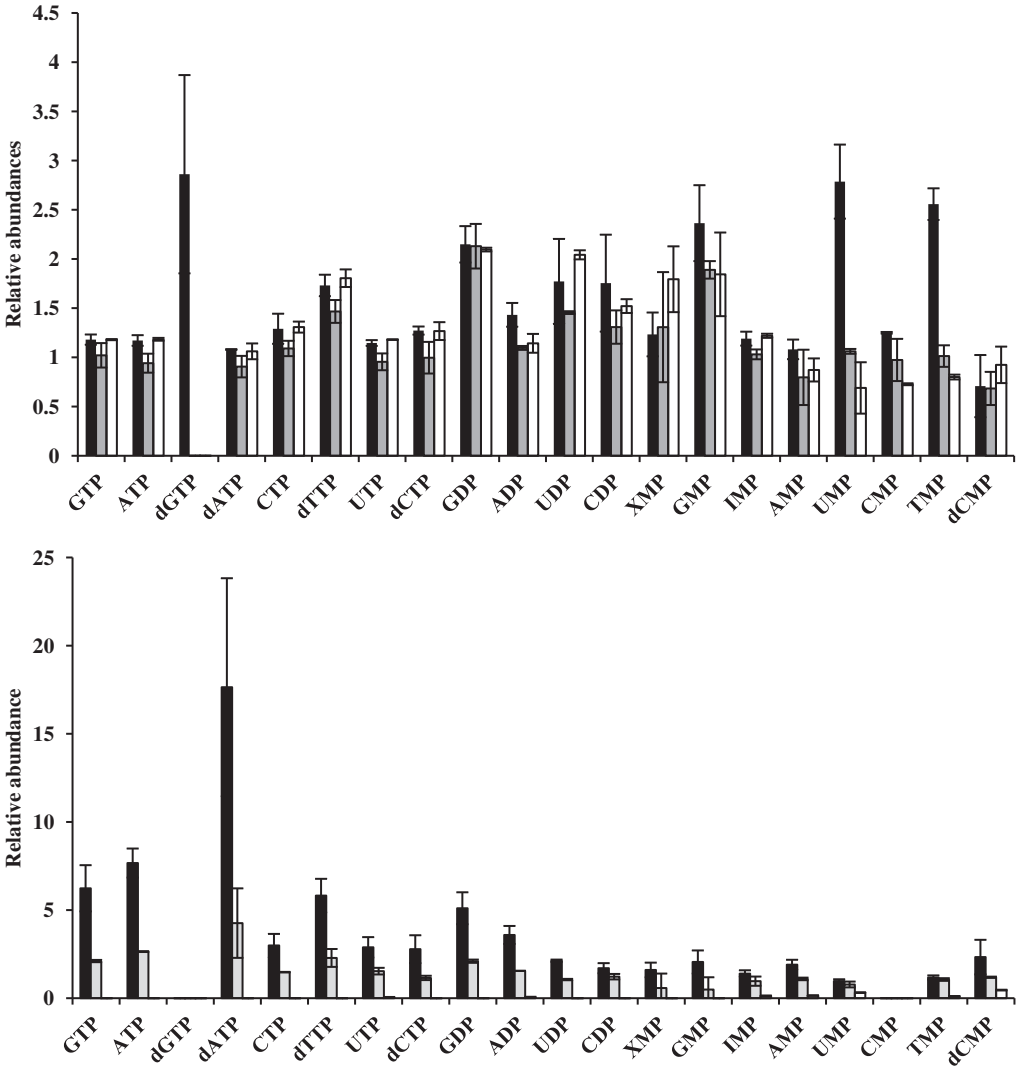


Fig. 3. Relative abundances of the metabolites detected in the *L. lactis* extracts after charcoal purification using different solvents for elution from the charcoal. Bar heights show the averages and standard deviations of duplicate samples. The bars show the differences in metabolite levels related to the different eluents used for elution of the metabolites from the charcoal relative to the amounts obtained when ethanol was used as organic modifier (relative abundance of 1). Legend for bars in upper graph: black, 50% acetonitrile and 2% ammonia in water; gray, 50% isopropanol and 2% ammonia in water; white, 50% methanol and 2% ammonia in water. Legend for bars in lower graph: black, 50% acetonitrile containing 30 mM DBAA; gray, 50% isopropanol containing 30 mM DBAA; horizontal lines, 50% methanol containing 30 mM DBAA.

pair reagent was not sufficiently strong for elution of most of the nucleotides from the charcoal in a quantitative manner, indicating that the high pH was more important than the nature of the organic modifier. Fully 50% less ATP was eluted from the charcoal by using an acetonitrile/DBAA mixture when compared with the amount of ATP eluted when using the acetonitrile/ammonia mixture.

Method validation and assessment of matrix effects

Due to the large difference in the intracellular concentration, the validation was performed at different concentration levels to assess the variation at relevant concentrations. Due to the lack of commercially available isotope-labeled analogs for all of the nucleotides, only those available were used and, thus, the results are divided into two groups: metabolites measured with IS and those measured without IS. Table 2 summarizes the validation data from the measurement of the lowest externally added amounts of a particular nucleotide that were measured with acceptable accuracy and precision. The bolded values, in the columns with the added amounts, represent the levels that fall into the accepted range of ±20% deviation, whereas the shown accuracy and precision data are from the measurement of the underlined values.

As expected, better recovery/accuracy, precision, and R^2 were obtained for the compounds quantitated using stable isotope analogs as ISs. It is important to mention that in order to obtain good linearity and acceptable precision and accuracy for all of the levels in the standard addition experiment, it was necessary to add IS at a concentration equal to or higher then the concentration of the compound detected in the sample. For example, ATP, UTP, and GTP are compounds present in higher amounts in the cell, and the externally added amounts of these compounds brought the concentration into a range where the IS was not capable of correcting for the variations unless it was present in a sufficient amount. When the amounts of the [^{13}C]ATP, [^{15}N]UTP, and [^{13}C]GTP were increased 10, 25, and 12 times, respectively, the linearity was improved from 0.966 to 0.992 for ATP, from 0.973 to 0.992 for UTP, and from 0.992 to 0.995 for GTP. This suggests that the IS should be added in similar concentrations as the target compounds, although this will be very costly (~U.S. $250/sample). An alterna-

tive could be to use intracellular metabolites extracted from cells grown on ^{13}C-labeled glucose, as published previously [20,33]. In our setup with the DSPE, we believe that it will not be an option in this analysis due to the additional impurities that will be introduced into the samples and will further affect the measurement.

Only UDP–glucose had accuracy and precision outside the allowed range of ±20% for all of the levels. The reason for the inaccurate measurement of UDP–glucose was assumed to be the high concentration of this compound present in the cell along with the externally added amounts in the standard addition experiment, which brought the concentration levels into the nonlinearity range. The latter, combined with the absence of isotope-labeled IS for this compound, resulted in the validation failing. To solve this problem, one could dilute the samples, which would require the sample to be analyzed twice due to the low concentration of some of the nucleotides.

Second validation approach was preformed by spiking the growth medium in concentrations ranging from 0 to 6 μg/ml before charcoal purification (Table 3). For the compounds for which IS was available, the measurements of the samples prepared by spiking the medium before charcoal purification showed linearity across the entire range with acceptable accuracy (80–120%) and precision (relative standard deviation [RSD] of ±20%). Due to the lower concentration of the nucleotides in the spiking samples, the IS was able to correct for both the matrix effects and the losses during the sample preparation. Regarding the nucleotides for which IS was not added, the linearity was in the range between 0.1 and 1.2 μg/ml in most of the cases. UDP–glucose also showed acceptable accuracy and precision in the range between 0.1 and 1.2 μg/ml, which adds to the previous assumption that the validation with the standard addition experiment failed due to the concentrations exceeding the linearity range.

In addition, the matrix effects were evaluated from the neat standards and the samples prepared by spiking the medium with standards and IS after charcoal purification. The reason for choosing the medium is due to the sample preparation procedure, where the cells are kept in the growth medium during the extraction and some of the medium components can interfere with the targets of interest.

The results are given in Table 4.

The classic T-piece infusion test [38] shown in Fig. 4 shows the drops in the signal intensity of ATP (m/z 505.9 > 407.9) in the chromatograms, indicating the presence of interfering compounds causing ion suppression. The strongest suppression was observed at the beginning of the gradient, presumably caused by salts. Although the nucleotides elute later in the method, they can still be affected by some impurities present in the samples due to the contamination of the ion source after several injections. In addition, Fig. 4 shows the effect on the signal intensity of ATP caused by injection of *L. lactis* sample purified with charcoal. As expected, more suppression is observed in this case, which can be assigned to both the medium and all of the other extracted metabolites from the cell during the sample preparation as well as self-suppression coming from the coeluting nucleotides.

Comparison with legacy data

The measured intracellular concentrations are given in Table 2. As mentioned in Materials and Methods, the intracellular concentrations were estimated assuming an intracellular volume of 1.67 ml for 1 g (dry weight) of cells [36]. Fig. 5 shows the comparison of the determined intracellular concentration by the ion-pair LC–MS/MS method with the data obtained using ^{33}P labeling followed by TLC.

Table 3
Repeatability and reproducibility of the measurements done by spiking the medium before charcoal purification.

		R^2	Repeatability (%)[a,b]										Reproducibility (%)									
			0.1 μg/ml		0.3 μg/ml		0.6 μg/ml		1.2 μg/ml		6 μg/ml		0.1 μg/ml		0.3 μg/ml		0.6 μg/ml		1.2 μg/ml		6 μg/ml	
	Compound		Acc	RSD	Acc	RSD	Acc	RSD	Acc	RSD	Acc	RSD	Acc	RSD	Acc	RSD	Acc	RSD	Acc	RSD	Acc	RSD
With IS	AMP	0.995	97	7	98	4	94	4	104	6	100	13	99	8	98	2	93	6	105	9	100	10
	ATP	0.994	105	7	94	12	99	5	102	8	101	11	105	10	99	13	99	5	103	9	102	11
	CMP	0.994	104	11	89	7	93	9	106	8	100	7	105	10	88	7	95	10	107	7	100	6
	CTP	0.996	100	11	90	8	87	11	89	4	103	15	99	10	91	8	83	10	90	3	100	12
	dATP	0.994	99	13	91	14	85	7	95	5	102	10	102	13	90	13	83	8	95	5	101	9
	dCMP	0.998	102	12	94	5	96	8	103	5	100	5	102	12	95	5	98	6	102	4	100	4
	dCTP	0.995	98	11	92	15	97	10	100	8	102	4	102	15	89	14	96	12	99	8	102	4
	dGMP	0.991	110	10	90	11	86	5	92	6	103	10	111	11	96	14	88	6	93	7	102	8
	dGTP	0.996	110	10	90	11	86	5	92	6	103	10	111	11	96	14	88	6	93	7	102	8
	dTMP	0.998	102	7	92	7	95	10	104	7	100	5	102	9	91	7	95	9	105	6	100	4
	GMP	0.991	116	19	81	19	88	6	96	5	102	6	108	16	85	18	89	6	96	5	101	5
	GTP	0.992	94	12	90	18	84	11	84	11	106	5	92	19	82	15	90	11	90	11	107	5
	TTP	0.992	97	11	93	5	88	6	96	4	102	9	98	11	93	5	87	6	97	4	102	9
	UMP	0.995	97	8	93	9	88	6	104	5	103	5	97	8	93	8	88	5	104	5	100	4
	UTP	0.992	90	16	94	8	105	8	95	7	96	7	91	15	95	10	103	12	101	7	94	20
Without IS	ADP	0.999	105	16	96	10	118	6	104	1	99	6	101	18	94	9	101	6	104	2	99	5
	cAMP	0.990	97	5	105	6	93	20	108	2	–	–	104	7	106	5	88	18	108	5	–	–
	cdiGMP	0.989	106	15	98	11	92	9	104	7	–	–	104	13	96	12	93	8	103	5	–	–
	CDP	0.943	–	–	95	15	106	7	98	7	–	–	–	–	95	13	108	9	97	6	–	–
	cGMP	0.988	103	8	99	10	97	18	103	2	–	–	102	8	99	9	91	18	103	2	–	–
	dUMP	0.990	108	12	96	5	98	8	114	5	–	–	92	12	94	5	96	6	115	4	–	–
	IMP	0.989	92	14	103	17	112	18	–		–	–	94	17	88	16	102	17	–	–	–	–
	OMP	0.973	91	17	107	18	103	12	97	9	–	–	96	20	103	17	101	11	98	8	–	–
	UDP-Glc	0.992	102	10	100	11	104	12	101	5	–	–	101	9	99	9	98	10	100	4	–	–
	UDP	0.970	99	16	97	14	97	14	102	7	–	–	100	18	96	14	97	12	102	7	–	–
	XMP	0.993	103	11	100	14	95	8	102	4	–	–	102	9	99	11	95	7	102	3	–	–
	ZMP	0.997	–	–	110	16	101	15	102	6	–	–	–	–	109	15	95	14	102	8	–	–
	GDP	0.988	–	–	95	8	99	9	105	5	–	–	–	–	95	12	98	8	107	10	–	–

[a] The intraday repeatability and interday reproducibility were determined by analyzing triplicates of five different concentration levels at 3 different days.
[b] The accuracy was determined as 100 − [(added-measured)/measured × 100].

Table 4
Matrix effects from the medium expressed as the ratio of the slope of the calibration curve constructed from standards spiked in the charcoal-purified medium, before the analysis, to the standards prepared in the neat solution multiplied by 100.

Nucleotide	Slope from the linear regression analysis		Matrix effects[a] (%)
	Neat standards	Standards spiked in the purified medium before injection into the LC–MS	
c-diGMP	12.8	8.0	63
UDP-Glc	109	80.9	74
GTP	53.8	37.9	71
ATP	49.7	37.0	74
dGTP	10.3	5.6	55
dATP	262	155	59
UTP	305	90.7	30
CTP	201	105	53
dTTP	338	133	39
dCTP	25.4	15.5	61
GDP	167	142	85
ADP	49.1	26.4	54
UDP	48.7	41.7	86
CDP	48.7	35.1	72
OMP	57.9	46.5	80
XMP	61.6	30.8	50
GMP	44.4	30.7	69
IMP	36.0	23.8	66
AMP	117	71.5	61
dGMP	57.4	22.5	39
cGMP	204	130	63
ZMP	15.4	7.9	51
cAMP	375	201	54
UMP	112	112	100
CMP	47.3	32.9	70
dTMP	103	51.7	50
dUMP	57.2	32.3	56
dCMP	63.6	57.5	90

Note: For abbreviations, see note in Table 1. The experimental details are given in Materials and Methods.
[a] 100% corresponds to no matrix effect.

Two different quantification approaches were used: (i) standard addition and (ii) external calibration by using calibration standards prepared by spiking the medium before charcoal purification. For the compounds for which IS was used, similar concentrations were obtained independent of the quantification approach. This once more confirms that the IS is capable of correcting the losses and the matrix effects during the analysis if present in a sufficient amount, as mentioned before. Regarding the compounds for which quantification was done without the addition of IS, different concentrations were obtained depending on the quantification approach used. The differences could be explained by the different matrix effects between the samples and the standard in the external calibration experiment and the absence of the IS in general. However, as has been previously published by Liu and coworkers [40], even the addition of isotope-labeled analog will not always improve the accuracy.

For most of the nucleotides for which IS was added, no statistical difference (proved by two-tailed *t* test [data not shown]) was observed for the amounts found in *L. lactis* with standard addition versus quantification based on spiking the medium (Fig. 5). However for compounds with no isotope-labeled standard, big differences were observed; thus, standard addition is essential for a correct quantification, whereas spiked medium will easily work for mutant or growth condition comparisons for all compounds quantified using isotope-labeled standards (at a similar analyte concentration).

The energy charge defined as (ATP + 0.5ADP)/(ATP + ADP + AMP) was determined from the quantified amounts of ATP, ADP, and AMP and was 0.97 ± 0.001, as reported previously in the literature [41].

The ratio between CTP and dCTP has been found to be between 5 and 10 [8], which is a general trait of many organisms, including mammalian cells [15]. To investigate whether the low CTP pool size was due to partial CTP starvation, *L. lactis* cultures were grown in SA medium with the addition of 200 μg/ml cytidine. This addition should boost the CTP pool of CTP-starved cells because cytidine is readily converted into CTP [15]. The results from the latter measurement showed the same trend. In addition, the extracts from the cells grown in a medium with and without the addition of cytidine were analyzed by the same UHPLC method now coupled to a QTOF instrument (high-resolution MS) to investigate the possible presence of coeluting compounds that might impair the analysis. The accurate mass UHPLC–MS measurement verified the dCTP/CTP ratio and showed that it was the two major ionizable compounds eluting at the specific time window, indicating no ion suppression. Finally, analysis of extracts from mamma-

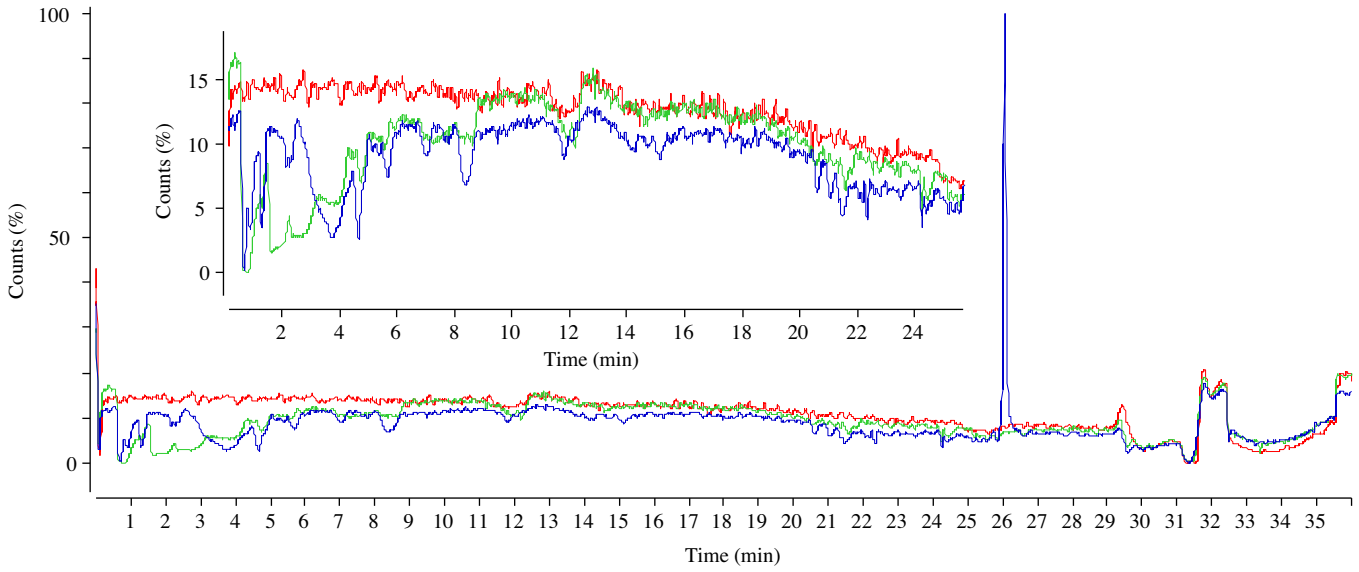


Fig. 4. Inspecting ion suppression by monitoring ATP signal from continuous infusion of ATP standard solution and injection of: i) mobile phase; ii) medium purified with charcoal; iii) and purified *L. lactis* sample. Legend: red signal, injection of mobile phase; blue signal, injection of *L. lactis* extract; green signal, injection of medium used for growing the cells purified with charcoal.

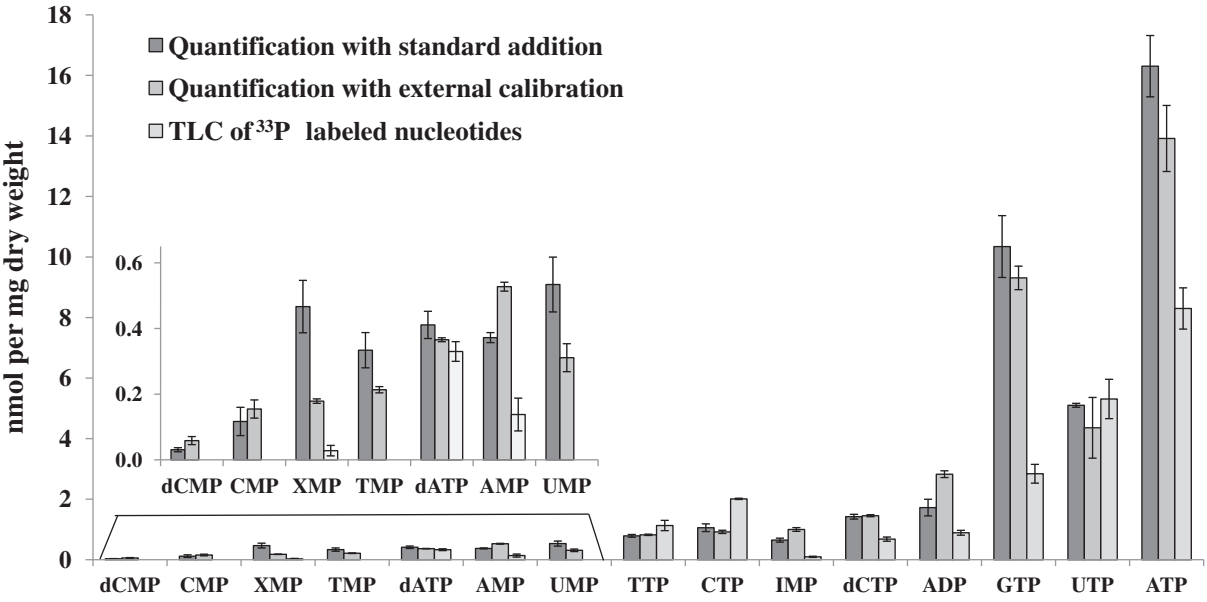


Fig. 5. Comparison of the measured intracellular concentrations by ion-pair LC–MS/MS using two quantification approaches with the concentrations determined using ^{33}P labeling followed by TLC (dCMP, CMP, TMP, and UMP were not measured by TLC) [9].

lian cells (Chinese hamster ovary [CHO] cells, saline quenched [42]) as well as *S. cerevisiae* (60% cold methanol, quenched) showed the expected CTP/dCTP ratio (results not shown), suggesting that the unexpected CTP/dCTP ratio is caused by cell physiology and not a technical error. When compared with the amounts of the nucleotides determined by TLC detection of radioactively ^{33}P-labeled nucleotides by Jendresen and coworkers [9] (Fig. 5), higher concentrations were measured with the LC–MS/MS method for most of the compounds. Because both LC–MS/MS calibration methods gave the same results, this supports the results obtained here and substantiates that the cells may have been in different physiological states in the two experimental setups. We are currently pursuing an explanation for these issues.

Conclusion

An analytical method for 28 nucleotides in *L. lactis* has been established. The method can, via charcoal purification, handle extracts where the fragile cells leaks during the quenching. In addition, separation of isobaric and most of the compounds that differ in one mass unit (e.g., dTMP, dUMP, dCMP) was achieved either by MS/MS (specific fragments) or chromatographically. The combination of ion-pair LC–MS/MS together with charcoal sample preparation proved to be suitable for simultaneous analysis of mono-, di-, and triphosphorylated deoxy- and ribonucleotides and partly sugar nucleotides. The method was successfully applied for analysis of nucleotide pools in *L. lactis* but also showed that for the standard addition experiment the isotope-labeled standards needed to be added in similar or higher concentrations as the analytes.

Acknowledgments

This work was supported by the Danish Research Agency for Technology and Production (grant 09-064967). The authors acknowledge the Technical University of Denmark Fermentation Platform at the Department of Systems Biology for providing the Agilent 6460 Triple Quadrupole LC/MS instrument used in this study. We thank Agilent Technologies for the Thought Leader Donation of the 6550 QTOF instrument.

References

[1] D. Kalia, G. Merey, S. Nakayama, Y. Zheng, J. Zhou, Y. Luo, M. Guo, B.T. Roembke, H.O. Sintim, Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria, and implications in pathogenesis, Chem. Soc. Rev. 42 (2013) 305–341.
[2] C.I. Pogson, Guanine nucleotides and their significance in biochemical processes, Am. J. Clin. Nutr. 27 (1974) 380–402.
[3] M. Conti, Phosphodiesterases and cyclic nucleotide signaling in endocrine cells, Mol. Endocrinol. 14 (2000) 1317–1327.
[4] U.B. Kaupp, R. Seifert, Cyclic nucleotide gated ion channels, Physiol. Rev. 82 (2002) 769–824.
[5] A.B. Canelas, C. Ras, A. ten Pierick, J.C. van Dam, J.J. Heijnen, W.M. van Gulik, Leakage-free rapid quenching technique for yeast metabolomics, Metabolomics 4 (2008) 226–239.
[6] S.G.V. Bôas, J.H. Pedersen, M. Åkesson, J. Smedsgaard, J. Nielsen, Global metabolite analysis of yeast: evaluation of sample preparation methods, Yeast 22 (2005) 1155–1169.
[7] C.J. Bolten, P. Kiefer, F. Letisse, J.C. Portais, C. Wittmann, Sampling for metabolome analysis of microorganisms, Anal. Chem. 79 (2007) 3843–3849.
[8] J. Martinussen, S. Wadskov-Hansen, K. Hammer, Two nucleoside uptake systems in *Lactococcus lactis*: competition between purine nucleosides and cytidine allows for modulation of intracellular nucleotide pools, J. Bacteriol. 185 (2003) 1503–1508.
[9] C.B. Jendresen, M. Kilstrup, J. Martinussen, A simplified method for rapid quantification of intracellular nucleoside triphosphates by one-dimensional thin-layer chromatography, Anal. Biochem. 409 (2011) 249–259.
[10] D. Mutavdzič Pavlovič, S. Babič, D. Dolar, D. Ašperger, K. Košutić, A.J. Horvat, M. Kastelan-Macan, Development and optimization of the SPE procedure for determination of pharmaceuticals in water samples by HPLC–diode array detection, J. Sep. Sci. 33 (2010) 258–267.
[11] A.R. Fontana, A. Camargo, L.D. Martinez, J.C. Altamirano, Dispersive solid-phase extraction as a simplified clean-up technique for biological sample extracts: determination of polybrominated diphenyl ethers by gas chromatography–tandem mass spectrometry, J. Chromatogr. A 1218 (2011) 2490–2496.
[12] S. Yang, R.E. Synovec, M.G. Kalyuzhnaya, M.E. Lidstrom, Development of a solid phase extraction protocol coupled with liquid chromatography mass spectrometry to analyze central carbon metabolites in lake sediment microcosms, J. Sep. Sci. 34 (2011) 3597–3605.
[13] M. Anastassiades, S.J. Lehotay, D. Stajnbaher, F.J. Schenck, Fast and easy multiresidue method employing acetonitrile extraction/partitioning and "dispersive solid-phase extraction" for the determination of pesticide residues in produce, J. AOAC Int. 86 (2003) 412–431.
[14] M. Cashel, R.A. Lazzarini, B. Kalbacher, An improved method for thin-layer chromatography of nucleotide mixtures containing ^{32}P-labeled orthophosphate, J. Chromatogr. 40 (1969) 103–109.
[15] C.M. Jørgensen, K. Hammer, P.R. Jensen, J. Martinussen, Expression of the *pyrG* gene determines the pool sizes of CTP and dCTP in *Lactococcus lactis*, Eur. J. Biochem. 271 (2004) 2438–2445.
[16] J. Hua, K.-L. Huang, A reversed phase HPLC method for the analysis of nucleotides to determine 5′-PDE enzyme activity, Bull. Chem. Soc. Ethiop. 24 (2010) 167–174.

[17] E. Witters, W. Van Dongen, E.L. Esmans, H.A. Van Onckelen, Ion-pair liquid chromatography–electrospray mass spectrometry for the analysis of cyclic nucleotides, J. Chromatogr. B 694 (1997) 55–63.
[18] M. Cichna, M. Raab, H. Daxecker, A. Griesmacher, M.M. Müller, P. Markl, Determination of fifteen nucleotides in cultured human mononuclear blood and umbilical vein endothelial cells by solvent generated ion-pair chromatography, J. Chromatogr. B 787 (2003) 381–391.
[19] G. Haink, A. Deussen, Liquid chromatography method for the analysis of adenosine compounds, J. Chromatogr. B 784 (2003) 189–193.
[20] R.M. Seifar, C. Ras, J.C. van Dam, W.M. van Gulik, J.J. Heijnen, W.A. van Winden, Simultaneous quantification of free nucleotides in complex biological samples using ion pair reversed phase liquid chromatography isotope dilution tandem mass spectrometry, Anal. Biochem. 388 (2009) 213–219.
[21] L. Coulier, R. Bas, S. Jespersen, E. Verheij, M.J. van der Werf, T. Hankemeier, Simultaneous quantitative analysis of metabolites using ion-pair liquid chromatography–electrospray ionization mass spectrometry, Anal. Chem. 78 (2006) 6573–6582.
[22] T. Van Damme, Y. Zhang, F.C. Lynen, P. Sandra, Determination of cyclic guanosine and cyclic adenosine monophosphate (cGMP and cAMP) in human plasma and animal tissues by solid phase extraction on silica and liquid chromatography–triple quadrupole mass spectrometry, J. Chromatogr. B 909 (2012) 14–21.
[23] K.Y. Beste, H. Burhenne, V. Kaever, J.-P. Stasch, R. Seifert, Nucleotidyl cyclase activity of soluble guanylyl cyclase $\alpha_1\beta_1$, Biochemistry 51 (2012) 194–204.
[24] C. Crauste, I. Lefebvre, M. Hovaneissian, J.Y. Puy, B. Roy, S. Peyrottes, S. Cohen, J. Guitton, C. Dumontet, C. Perigaud, Development of a sensitive and selective LC–MS/MS method for the simultaneous determination of intracellular 1-β-D-arabinofuranosylcytosine triphosphate (araCTP), cytidine triphosphate (CTP), and deoxycytidine triphosphate (dCTP) in a human follicular lymphoma cell line, J. Chromatogr. B 877 (2009) 1417–1425.
[25] T. Kind, G. Wohlgemuth, D.Y. Lee, Y. Lu, M. Palazoglu, S. Shahbaz, O. Fiehn, FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry, Anal. Chem. 81 (2009) 10038–10048.
[26] K. Inoue, R. Obara, T. Hino, H. Oka, Development and application of an HILIC–MS/MS method for the quantitation of nucleotides in infant formula, J. Agric. Food Chem. 58 (2010) 9918–9924.
[27] K. Inoue, R. Obara, T. Akiba, T. Hino, H. Oka, Determination of nucleotides in infant formula by ion-exchange liquid chromatography, J. Agric. Food Chem. 56 (2008) 6863–6867.
[28] N.B. Bennette, J.F. Eng, G.C. Dismukes, An LC–MS-based chemical and analytical method for targeted metabolite quantification in the model *Cyanobacterium synechococcus* sp. PCC 7002, Anal. Chem. 83 (2011) 3808–3816.
[29] B. Luo, K. Groenke, R. Takors, C. Wandrey, M. Oldiges, Simultaneous determination of multiple intracellular metabolites in glycolysis, pentose phosphate pathway, and tricarboxylic acid cycle by liquid chromatography–mass spectrometry, J. Chromatogr. A 1147 (2007) 153–164.
[30] M. Holčapek, K. Volná, P. Jandera, L. Kolářová, K. Lemr, M. Exner, A. Církva, Effects of ion-pairing reagents on the electrospray signal suppression of sulphonated dyes and intermediates, J. Mass Spectrom. 39 (2004) 43–50.
[31] H. Trufelli, P. Palma, G. Famiglini, A. Cappiello, An overview of matrix effects in liquid chromatography–mass spectrometry, Mass Spectrom. Rev. 30 (2011) 491–509.
[32] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC–MS/MS, Anal. Chem. 75 (2003) 3019–3030.
[33] L. Wu, M.R. Mashego, J.C. van Dam, A.M. Proell, J.L. Vinke, C. Ras, W.A. van Winden, W.M. van Gulik, J.J. Heijnen, Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ^{13}C-labeled cell extracts as internal standards, Anal. Biochem. 336 (2005) 164–171.
[34] B.E. Terzaghi, W.E. Sandine, Improved medium for lactic streptococci and their bacteriophages, Appl. Microbiol. 29 (1975) 807–813.
[35] P.R. Jensen, K. Hammer, Minimal requirements for exponential growth of *Lactococcus lactis*, Appl. Environ. Microbiol. 59 (1993) 4363–4366.
[36] J. Thompson, Characteristics and energy requirements of an α-aminoisobutyric acid transport system in *Streptococcus lactis*, J. Bacteriol. 127 (1976) 719–730.
[37] D. Tsikas, A proposal for comparing methods of quantitative analysis of endogenous compounds in biological systems by using the relative lower limit of quantification (rLLOQ), J. Chromatogr. B 877 (2009) 2244–2251.
[38] T.M. Annesley, Ion suppression in mass spectrometry, Clin. Chem. 49 (2003) 1041–1044.
[39] J.M. Buescher, S. Moco, U. Sauer, N. Zamboni, Ultra high performance liquid chromatography–tandem mass spectrometry method for fast and robust quantification of anionic and aromatic metabolites, Anal. Chem. 82 (2010) 4403–4412.
[40] G. Liu, Q.C. Ji, M.E. Arnold, Identifying, evaluating, and controlling bioanalytical risks resulting from nonuniform matrix ion suppression/enhancement and nonlinear liquid chromatography–mass spectrometry assay response, Anal. Chem. 82 (2010) 9671–9677.
[41] M. Faijes, A.E. Mars, E.J. Smid, Comparison of quenching and extraction methodologies for metabolome analysis of *Lactobacillus plantarum*, Microb. Cell Fact. 6 (2007) 27.
[42] S. Dietmair, N.E. Timmins, P.P. Gray, L.K. Nielsen, J.O. Kromer, Towards quantitative metabolomics of mammalian cells: development of a metabolite extraction protocol, Anal. Chem. 404 (2010) 155–164.

6.2 Paper 2 - Biofilm formation and antibiotic production in *Ruegeria mobilis* are influenced by intracellular concentrations of cyclic dimeric guanosin monophosphate

Paul W. D'Alvise; **Olivera Magdenoska**; Jette Melchiorsen; Kristian Fog Nielsen; Lone Gram.

Paper published in Environmental Microbiology 2014

environmental microbiology

Environmental Microbiology (2013) doi:10.1111/1462-2920.12265

Biofilm formation and antibiotic production in *Ruegeria mobilis* are influenced by intracellular concentrations of cyclic dimeric guanosinmonophosphate

Paul W. D'Alvise,* Olivera Magdenoska, Jette Melchiorsen, Kristian F. Nielsen and Lone Gram

Department of Systems Biology, Technical University of Denmark, Søltofts Plads, Bldg. 221, DK-2800 Kgs. Lyngby, Denmark.

Summary

In many species of the marine *Roseobacter* clade, periods of attached life, in association with phytoplankton or particles, are interspersed with planktonic phases. The purpose of this study was to determine whether shifts between motile and sessile life in the globally abundant *Roseobacter* clade species *Ruegeria mobilis* are associated with intracellular concentrations of the signal compound cyclic dimeric guanosinmonophosphate (c-di-GMP), which in bacteria regulates transitions between motile and sessile life stages. Genes for diguanylate cyclases and phosphodiesterases, which are involved in c-di-GMP signalling, were found in the genome of *R. mobilis* strain F1926. Ion pair chromatography-tandem mass spectrometry revealed 20-fold higher c-di-GMP concentrations per cell in biofilm-containing cultures than in planktonic cells. An introduced diguanylate cyclase gene increased c-di-GMP and enhanced biofilm formation and production of the potent antibiotic tropodithietic acid (TDA). An introduced phosphodiesterase gene decreased c-di-GMP and reduced biofilm formation and TDA production. *tda*C, a key gene for TDA biosynthesis, was expressed only in attached or biofilm-forming cells, and expression was induced immediately after initial attachment. In conclusion, c-di-GMP signalling controls biofilm formation and biofilm-associated traits in *R. mobilis* and, as suggested by presence of GGDEF and EAL domain protein genes, also in other *Roseobacter* clade species.

Received 10 May, 2013; revised 19 August, 2013; accepted 26 August, 2013. *For correspondence. E-mail pdal@bio.dtu.dk; Tel. (+45) 45255501; Fax (+45) 45884922 .

Introduction

The *Roseobacter* clade (*Alphaproteobacteria*) accounts for a significant part of the microbiota in the oceans, especially in coastal zones and surface waters (Gonzalez and Moran, 1997; Buchan *et al*., 2005; Brinkhoff *et al*., 2008; Newton *et al*., 2010; Wietz *et al*., 2010). *Roseobacter* clade species are metabolically and ecologically diverse, comprising aerobic anoxygenic phototrophs, sulphur metabolizers, carbon monoxide oxidizers and degraders of aromatic compounds (Shiba *et al*., 1979; Sorokin and Lysenko, 1993; Moran and Hodson, 1994; Buchan *et al*., 2001; Allgaier *et al*., 2003; Moran *et al*., 2003; 2004). However, most species of the *Roseobacter* clade are classified as ecological generalists (Moran *et al*., 2004; Newton *et al*., 2010). Abundance and activity of many *Roseobacter* clade members are correlated with phytoplankton population densities, and one prominent ability of many *Roseobacter* clade members is the conversion of the phytoplankton osmolyte dimethylsulfoniopropionate to the volatile dimethyl sulphide, which influences the local and global climate (Charlson *et al*., 1987; Gonzalez and Moran, 1997; Gonzalez *et al*., 2000; Moran *et al*., 2003; Geng and Belas, 2010b).

Belas and colleagues (2009) noticed that many *Roseobacter* clade species have a 'biphasic swim-or-stick lifestyle' that enables their symbiosis with phytoplankton, and suggested that a central regulation mechanism coordinated the shift between planktonic and attached phenotype. Accumulating evidence indicates that bis-(3′-5′)-cyclic dimeric guanosinmonophosphate (c-di-GMP) functions as a nearly universal second messenger in bacteria, regulating transitions between planktonic and sedentary phases by controlling phenotypic features such as flagellar motility and production of extracellular polymeric substances (EPS) (Hengge, 2009; McDougald *et al*., 2012). Above that, c-di-GMP regulates important functions that are associated with either one of the lifestyles, such as virulence or antibiotic production (Schmidt *et al*., 2005; Cotter and Stibitz, 2007; Tamayo *et al*., 2007). The intracellular pool of c-di-GMP is balanced by diguanylate cyclases (GGDEF-domain proteins) that synthesize the compound and by specific phosphodiesterases (EAL-domain proteins) that degrade it (Ausmees *et al*., 2001;

Ryjenkov *et al*., 2005; Gjermansen *et al*., 2006). The activity of these antagonistic enzymes is controlled by sensory domains or proteins that allow external or internal stimuli to act on the intracellular pool of c-di-GMP and thus influence the decision between sessile and motile life. Here, we hypothesized that the transition between planktonic and attached lifestyle in *Roseobacter* clade species is induced by intracellular c-di-GMP levels. To test this, we introduced the plasmids pYedQ and pYhjH that have been used as tools to demonstrate that c-di-GMP signalling regulated *Pseudomonas putida* biofilm formation and dispersal (Gjermansen *et al*., 2006). The plasmids contain either one of the *Escherichia coli* genes *yed*Q and *yhj*H, which encode a diuguanylate cyclase that synthesizes c-di-GMP and a c-di-GMP-degrading phosphodiesterase respectively. Intracellular levels of c-di-GMP in *Ruegeria mobilis* F1926 wild-type and plasmid-carrying mutants were assessed using ion pair liquid chromatography-tandem mass spectrometry (Ion-Pair UHPLC-MS/MS).

The *Roseobacter* clade genera *Ruegeria* and *Phaeobacter* have been of particular interest due to their ability to form the potent antibacterial compound tropodithietic acid (TDA). *Phaeobacter* strains have been isolated from coastal zones, especially from biofilms in fish and invertebrate larvae cultures, and have been studied for their antibacterial activity and as probiotics for marine aquaculture (Ruiz-Ponte *et al*., 1998; Brinkhoff *et al*., 2004; Hjelm *et al*., 2004; Planas *et al*., 2006; Porsby *et al*., 2008; D'Alvise *et al*., 2010). *Ruegeria* (*Silicibacter*) sp. TM1040 is a symbiont of the dinoflagellate *Pfisteria piscicida* and was the first TDA-producing bacterium in which parts of the TDA biosynthetic pathway were elucidated (Alavi *et al*., 2001; Miller and Belas, 2006; Bruhn *et al*., 2007; Moran *et al*., 2007; Yi *et al*., 2007; Geng *et al*., 2008; Geng and Belas, 2010a). *R. mobilis* (*Ruegeria pelagia*) has been isolated in surface ocean waters in most climatic zones, except Arctic and Antarctic waters, and it occurs in both coastal zones and open oceans (Gram *et al*., 2010; Lai *et al*., 2010). Several strains of *R. mobilis*, which all produce TDA, have been isolated from microalgae cultures, highlighting their preference for phytoplankton (Porsby *et al*., 2008).

TDA production is influenced by culture conditions, and Bruhn and colleagues (2005) found that TDA production by *Phaeobacter* sp. 27-4 in Marine Broth (MB; Difco Laboratories, Detroit, MI, USA) only occurred under static growth conditions but not in shaken broth cultures. TDA production coincided with formation of a thick layer of biofilm at the air–liquid interface consisting of multicellular, star-shaped aggregates. To our knowledge, *Ruegeria* strains can only produce TDA in static liquid cultures, where a biofilm at the air–liquid interface is formed, whereas most *Phaeobacter* strains are able to produce TDA in shaken liquid cultures as well. This makes *Phaeobacter* sp. 27-4 an exemption among the TDA-producing *Phaeobacter* strains (Porsby *et al*., 2008). A later study, also using strain 27-4, demonstrated that attachment to an inert surface was affected by culture conditions as well (Bruhn *et al*., 2006). However, although the tight association between biofilm formation and TDA production strongly indicates that TDA is only produced by attached or biofilm-forming cells, this was never verified. In the present study, we demonstrate that c-di-GMP plays a major role in the transition from motile to sessile state in *R. mobilis*, and this led us to hypothesize that formation of TDA, which is associated with biofilms, could be correlated with c-di-GMP. Consequently, investigating whether TDA production was affected by altered c-di-GMP levels became a second aim of this study. Therefore, we measured the impact of changed c-di-GMP levels on antibacterial activity and TDA production and studied the expression of the *tda*C gene in attached and planktonic cells.

Results

Bioinformatic analysis

The genome of *R. mobilis* F1926 was Illumina-sequenced and assembled into 1065 contigs with a total length of about 4.5 Mb. One hundred ten contigs were larger than 10 kb and contained together 2.5 Mb. For comparison, the closest genome-sequenced relative (on 16S-rRNA-gene-level) of strain F1926, *Ruegeria* sp. TM1040 (Supporting Information Fig. S1), contains 4.2 Mb of genomic DNA, a megaplasmid of 0.8 Mb and a large plasmid of 0.1 Mb (Moran *et al*., 2007). Nine genes encoding diguanylate cyclases and c-di-GMP-specific phosphodiesterases were identified in *R. mobilis* F1926 based on Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) annotation, which was manually controlled by Pfam classification (Table 1). Six of these encoded proteins that contained both a GGDEF and an EAL domain. This could indicate that these enzymes have alternating c-di-GMP synthesizing or degrading activity; however, there are examples of proteins that contain both GGDEF and EAL domains, but act only as either diguanylate cyclase or phosphodiesterase, or have no c-di-GMP converting activity at all, but act as signalling proteins (Christen *et al*., 2005; Matilla *et al*., 2011; Newell *et al*., 2011).

c-di-GMP analysis by ion pair UHPLC-MS/MS

The plasmids pYedQ and pYhjH, as well as the respective vector controls pRK404A and pBBR1MCS-3, were introduced into *R. mobilis* F1926 with the aim of manipulating intracellular c-di-GMP concentrations, which was

Table 1. GGDEF- and EAL-domain protein genes identified in *Ruegeria mobilis* F1926.

Locus tag	Predicted gene product	Pfam domain	E-value (Pfam)	Length [aa]
K529_01085	diguanylate cyclase	GGDEF	4.1×10^{-29}	313
K529_06935	diguanylate cyclase/phosphodiesterase	GGDEF	2.5×10^{-35}	684
		EAL	5.4×10^{-66}	
K529_19462	diguanylate cyclase/phosphodiesterase	GGDEF	1.7×10^{-45}	700
		EAL	1.7×10^{-59}	
K529_09228	diguanylate cyclase	GGDEF	3.3×10^{-39}	342
K529_12495	diguanylate cyclase domain-containing protein	GGDEF	2.8×10^{-34}	674
		EAL	6.4×10^{-67}	
K529_13194	diguanylate cyclase/phosphodiesterase	GGDEF	6.9×10^{-6}	393
		EAL	7.6×10^{-57}	
K529_15333	diguanylate cyclase/phosphodiesterase	GGDEF	7.7×10^{-23}	496
		EAL	2.3×10^{-57}	
K529_15543	diguanylate cyclase/phosphodiesterase ammonium transporter	GGDEF	5.7×10^{-35}	904
		EAL	4.4×10^{-59}	
		Ammonium_transp	1.8×10^{-96}	
K529_20992	response regulator receiver modulated diguanylate cyclase	GGDEF	8.0×10^{-46}	467

subsequently verified by ion pair chromatography-MS/MS in extracts of 24 h old shaken and static MB cultures. The compound was detected in all extracts with the same retention time as an authentic standard. A chromatographic plot of a standard and a sample is given in the Supporting Information Fig. S2. Concentrations in the extracts ranged from 26 to 770 nM and were divided by the optical density (OD600) of the original culture to obtain a relative measure of c-di-GMP per cell (Fig. 1). Care was taken to break up the biofilms of the static cultures before sampling and OD measurement to allow comparison of cell density to shaken cultures. Static cultures of the wild type, where thick air liquid interface biofilms were observed, contained 20 times more c-di-GMP per cell than shaken cultures, where no biofilms were formed. Plasmid pYedQ increased c-di-GMP per cell to the 30-fold concentration of the vector control under shaken conditions ($P < 0.001$) and to the threefold concentration in static cultures ($P < 0.001$). In the static cultures of the mutant carrying plasmid pYhjH, c-di-GMP concentration per cell was reduced to about half of the respective vector control; however, this was not significant ($P > 0.05$), and under shaken conditions, no further reduction in c-di-GMP was found. Regardless of the introduced plasmids, cultivation conditions influenced c-di-GMP levels. For all strains, c-di-GMP concentrations per cell were higher in static cultures than in shaken cultures ($P < 0.001$).

Phenotypic effects of altered c-di-GMP levels

Shaken cultures of *R. mobilis* F1926 wild type were dominated by single cells, and about half of these were motile (Fig. 2A). Introduction of pYedQ, which increased c-di-GMP concentrations, caused disappearance of motile cells and increased formation of multicellular aggregates in shaken cultures (Fig. 2C). *R. mobilis* F1926 pYhjH, which contained less c-di-GMP, grew almost exclusively as single motile cells in shaken cultures (Fig. 2E). Static cultures of the wild type were dominated by biofilms consisting of multicellular, star-shaped aggregates (rosettes), but also, motile single cells were present (Fig. 2B). In static cultures of the pYedQ-carrying mutant, even thicker rosette-containing biofilms were formed, and no motile cells were observed (Fig. 2D). Introduction of pYhjH increased the proportion of motile cells and prevented formation of rosettes in static cultures (Fig. 2F). However, no difference in motility on population level could be detected in a motility-agar test (data not shown).

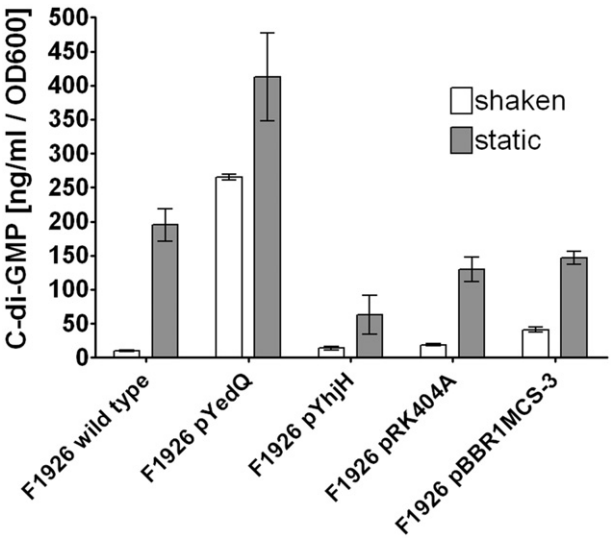


Fig. 1. Cyclic di-GMP concentrations in extracts of 24 h old shaken and static cultures of *Ruegeria mobilis* F1926 wild type, F1926 pYedQ, F1926 pYhjH and the respective vector control strains F1926 pRK404A and F1926 pBBR1MCS3, divided by OD600 of the cultures. Given values are averages of three independent replicates.

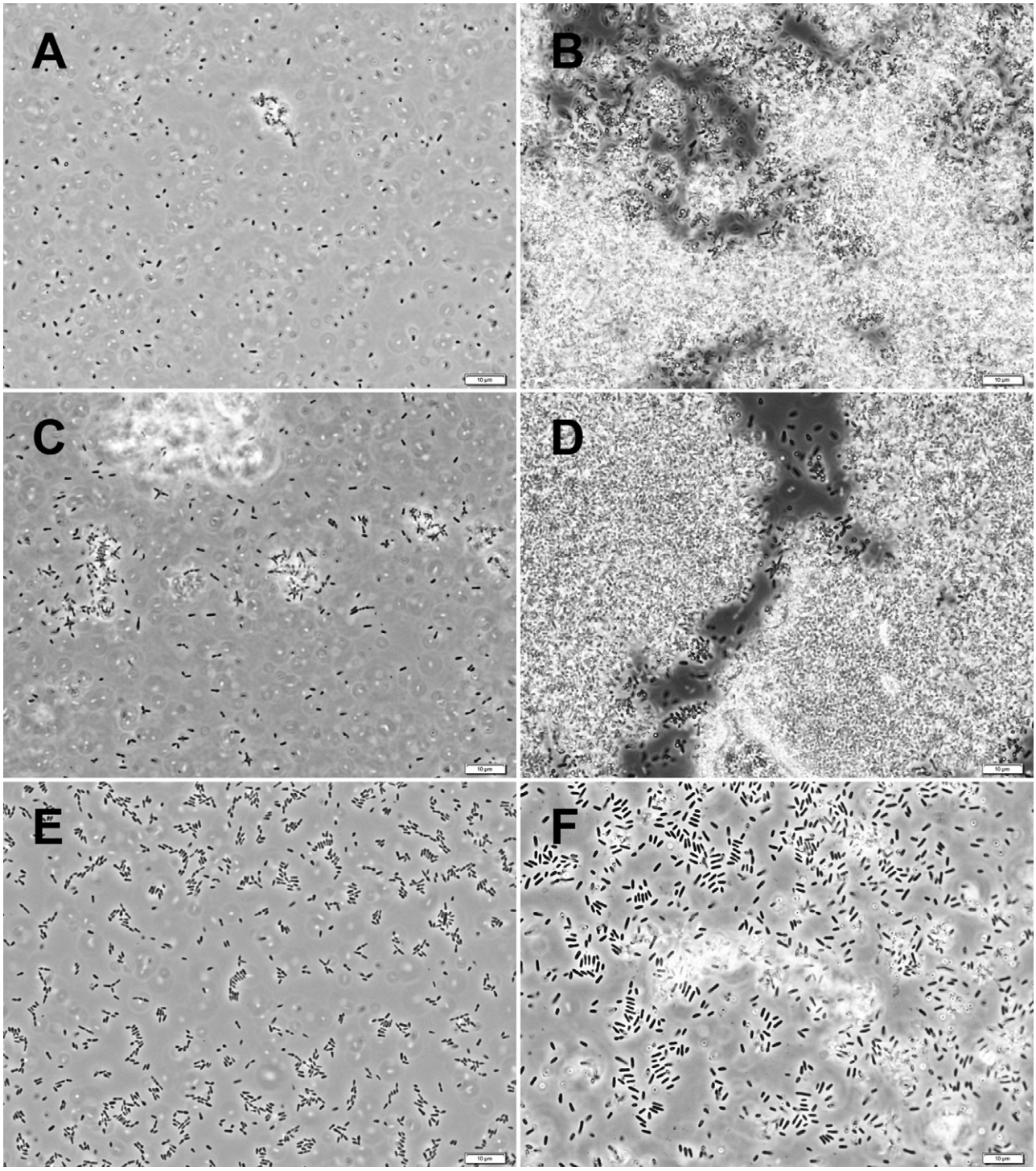


Fig. 2. Phase-contrast micrographs of shaken (A, C, E) and static (B, D, F) Marine Broth cultures of *Ruegeria mobilis* F1926 wild type (A, B), F1926 pYedQ (C, D) and F1926 pYhjH (E, F). All scale bars indicate 10 μm. The vector control strains were similar to the wild type (not shown).

Biofilm formation and attachment were assessed in the plasmid-carrying strains and wild type (Fig. 3). Biofilm formation (Fig. 3A) was significantly increased in F1926 pYedQ ($P < 0.001$), whereas in F1926 pYhjH, it was reduced as compared with the respective vector controls and the wild type ($P > 0.001$). Attachment of planktonic cells from shaken cultures was measured as stained biomass of cells attaching to polystyrene pegs within

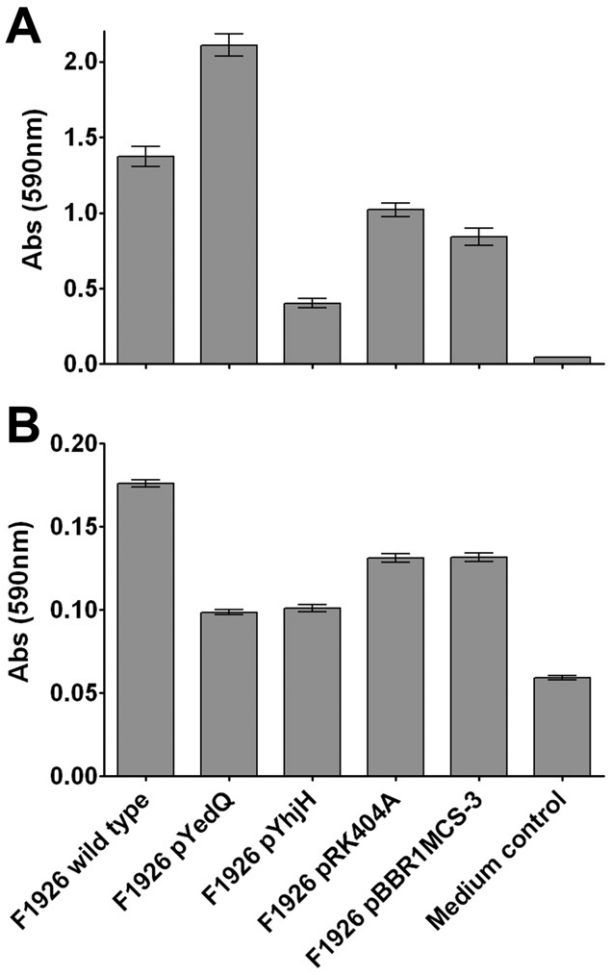


Fig. 3. Biofilm formation (A) and attachment (B) of *Ruegeria mobilis* F1926 wild type, F1926 pYedQ and F1926 pYhjH, and the respective vector control strains F1926 pRK404A and F1926 pBBR1MCS3, measured in crystal violet assays. Cells from shaken cultures were used in the attachment assay. Sterile medium (1.5 YTSS) was used as control. Given values are averages of two independent replicates.

1 min (Fig. 3B). The wild type attached most efficiently, whereas the attachment was reduced in both the pYhjH-carrying strain and the strain carrying pYedQ, indicating that both a decrease and an increase of intracellular c-di-GMP levels interfere with attachment. Although attachment was also reduced in the vector controls, which might be an effect of the tetracycline in the medium, the reduction of attachment conferred by the introduced GGDEF- and EAL-domain proteins was significant ($P < 0.001$).

The antibacterial effect of the plasmid-carrying mutants and the wild type was compared in agar-diffusion-inhibition assays, and TDA production was measured by ultra-performance liquid chromatography high-resolution mass spectrometry (UHPLC-TOFMS). In static cultures, the pYedQ-carrying mutant caused inhibition zones of equal or larger size than the wild type and vector controls, whereas the mutant carrying pYhjH was less inhibitory (Table 2). As opposed to the wild type, vector controls and the pYhjH-carrying mutant, the pYedQ-carrying mutant caused inhibition also in shaken cultures. As indicated by the inhibition test, TDA production was influenced by intracellular c-di-GMP concentrations (Fig. 4). Brown pigmentation of the cultures, which is indicative of TDA production (Bruhn *et al*., 2005), was correlated with inhibition and TDA production (not shown).

Expression of tda*C on single-cell level*

TDA production has been suggested to be regulated on community level by auto-induction (Geng and Belas, 2010a) or by quorum sensing (QS) (Berger *et al*., 2011). However, the observation that inhibitory activity was influenced by intracellular c-di-GMP levels, which can be different between attached and planktonic cells of the same community, suggested that TDA production could differ in cells within the same culture. This prompted us to study the expression of a key gene involved in TDA biosynthesis (*tda*C) on single-cell level by using a promoter-*gfp* fusion.

In shaken cultures, *R. mobilis* F1926 pPDA11 (P*tda*C::*gfp*) cells were predominantly planktonic (Fig. 5B), and approximately half the cells were motile, as described above for the wild type. *tda*C was not expressed, as indicated by lack of green fluorescence (Fig. 5A and B), correlating with a lack of antibacterial activity and TDA formation (Table 2, Fig 4). Yet a few small multicellular aggregates in which *tda*C was expressed were found (Fig. 5A and B). In contrast, major proportions of the cells from static cultures were situated in biofilms or multicellular aggregates and expressed the *tda*C gene (Fig. 5C and D). However, a part of the single cells in the samples from static cultures was not fluorescent, and a fraction of these was motile. No expression of *tda*C, as indicated by green fluorescence, was observed in motile cells. This indicated that *tda*C, a gene encoding a central enzyme in the TDA biosynthesis pathway, is expressed differently in adjacent cells within the same cultures, and

Table 2. Inhibition of *V. anguillarum* 90-11-287 by cell-free supernatants of *R. mobilis* F1926 wild type, F1926 pYedQ, F1926 pYhjH, F1926 pRK404A and F1926 pBBR1MCS-3 cultures, grown in shaken (200 r.p.m.) or static MB for 72 h at 25°C.

	Inhibition zone diameter without well diameter [mm] ± standard deviation	
Strain	Shaken cultures	Static cultures
R. mobilis F1926 wild type	0 ± 0	13.5 ± 0.7
R. mobilis F1926 pYedQ	1.5 ± 0.7	15 ± 0
R. mobilis F1926 pYhjH	0 ± 0	8 ± 1.4
R. mobilis F1926 pRK404A	0 ± 0	12 ± 0
R. mobilis F1926 pBBR1MCS-3	0 ± 0	12 ± 1,4

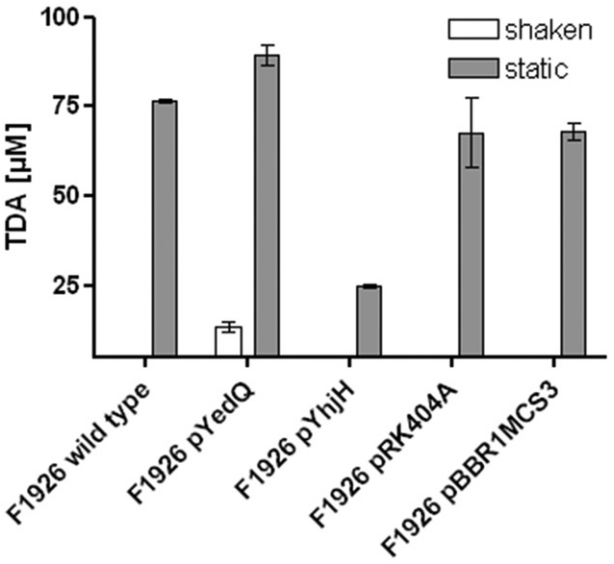


Fig. 4. Concentrations of TDA in shaken and static cultures of *Ruegeria mobilis* F1926 wild type, F1926 pYedQ, F1926 pYhjH and the respective vector control strains F1926 pRK404A and F1926 pBBR1MCS3, grown in shaken or static MB for 72 h at 25°C. Values are averages of two independent replicates.

that its expression coincides with attachment and is enhanced in biofilms or aggregates where high c-di-GMP levels are found.

Induction of tda*C expression in response to attachment*

When a sample of a shaken *R. mobilis* F1926 pPDA11 culture was prepared for microscopy, some of the non-motile single cells attached immediately to the cover glass. A time series of fluorescence and phase contrast micrographs showing newly attached cells on the cover glass was recorded (Fig. 6). *Gfp* expression was initiated in the newly attached cells, and green fluorescence could be observed already 10 min after preparation of the microscope slide. Maximal fluorescence of the initially attached cells was reached about 15–20 min after start of the experiment. Considering the time needed for *gfp* gene expression and maturation, this suggests that expression was initiated immediately when the cells attached. This instant initiation of *tda*C expression may indicate a regulatory connection between TDA production and initial attachment, which is likely accompanied by rising c-di-GMP levels.

Discussion

Attachment to surfaces and biofilm formation are characteristic features of many *Roseobacter* clade species, and a more comprehensive understanding of the transition between motile and sessile life stages in the *Roseobacter* clade is needed to understand carbon and nutrient cycling in the oceans (Slightom and Buchan, 2009). Also, production of the antibacterial substance TDA in *Phaeobacter* spp. has been associated with biofilms (Bruhn *et al*., 2005), and since TDA is a key component for their probiotic effect on fish larvae (D'Alvise *et al*., 2012), understanding the transition between motile and sessile stages may provide new perspectives on their application.

Many bacteria have two distinct lifestyles, a sessile or biofilm stage on a substrate or within a host, which is mostly characterized by increased metabolic activity and proliferation, and a mobile stage, where the cells are metabolically less active, disperse into the wider environment and persist until a new substrate or host is found. The transition between these two distinct states is in many bacteria controlled by a pool of intracellular c-di-GMP that exerts control on every level of regulation. C-di-GMP binds to transcriptional regulators and also exerts control at translational and allosteric levels, regulating a multitude of phenotypic traits, e.g. motility, EPS production, extracellular appendage formation, virulence or production of secondary metabolites (Hengge, 2009; McDougald *et al*., 2012). The results of our study suggest that c-di-GMP signalling plays a similar role in *R. mobilis*. Intracellular concentrations of c-di-GMP changed with cultivation conditions that favoured or prevented biofilm formation. Manipulation of c-di-GMP levels seemed to alter the proportions of motile and sessile cells under different culture conditions and affected expression of phenotypes that are associated with either planktonic or attached lifestyle. Increased levels of c-di-GMP promoted biofilm formation, whereas a decrease in c-di-GMP concentrations prevented formation of star-shaped aggregates and reduced TDA production. The finding that an introduced GGDEF domain protein increased biofilm formation and interfered with motility is consistent with previous studies (Ausmees *et al*., 2001; Simm *et al*., 2004; Gjermansen *et al*., 2006; Wolfe and Visick, 2008; Hengge, 2009; McDougald *et al*., 2012) and indicates a functional c-di-GMP signalling system in *R. mobilis*. This is confirmed by presence of genes encoding GGDEF and EAL domain proteins in the genome of *R. mobilis* F1926. Presence of similar genes annotated as diguanylyl cyclases and c-di-GMP-specific phosphodiesterases in the closely related strains *Ruegeria* sp. TM1040, *Ruegeria* sp. R11 and *Phaeobacter gallaeciensis* DSM17395, as well as in more remote *Roseobacter* clade species such as *Roseobacter litoralis* Och149, *Roseovarius* sp. 217, *Octadecabacter arcticus* 238 and *Dinoroseobacter shibae* DFL12 (Supporting Information Table S1), suggests that c-di-GMP signalling is a universal feature of the *Roseobacter* clade. We conclude that in the *Roseobacter* clade, in analogy with many other bacteria, intra and

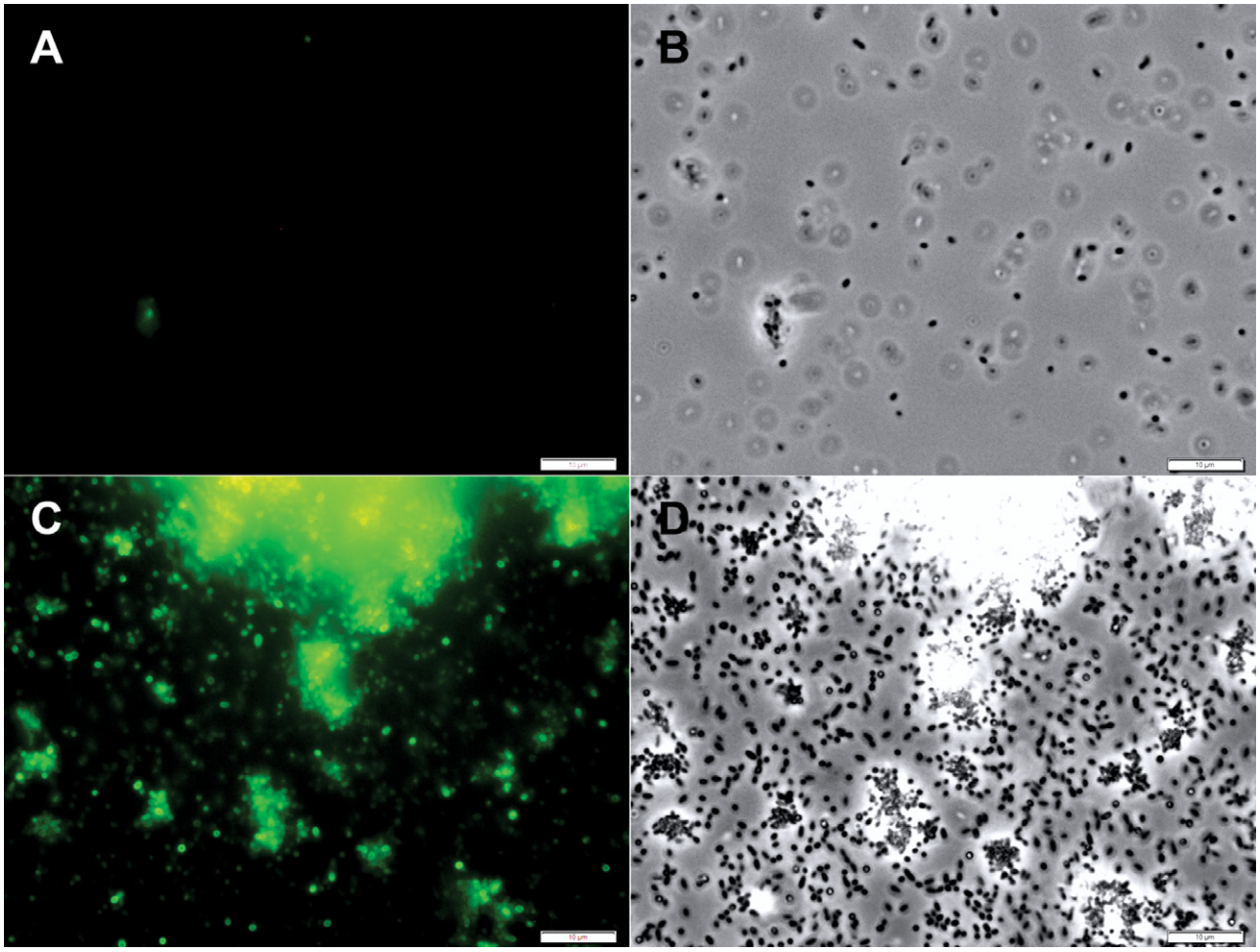


Fig. 5. Expression of *tda*C in *Ruegeria mobilis* F1926 grown in shaken (A, B) and static (C, D) MB cultures, monitored using a plasmid-bound reporter fusion of the *tda*C-promoter with a promoterless *gfp*-gene (pPDA11). Fluorescence (A, C) and phase contrast (B, D) micrographs were each recorded using the same settings.

extracellular cues are integrated via a c-di-GMP second messenger system and that expression of phenotypic traits specific for either planktonic or attached life is regulated in response to c-di-GMP concentrations. Belas and colleagues (2009) have introduced the term 'swim-or-stick switch' for the molecular mechanism that regulates transitions between motile and sessile stage in *Roseobacter* clade species. We think that the intracellular concentration of c-di-GMP is the swim-or-stick switch.

Recently, Zan and colleagues (2012) revealed that in *Ruegeria* sp. KLH11, a sponge symbiont, motility and biofilm formation are controlled by N-acyl homoserine lactone-based QS. Similarly, Sule and Belas (2012) found that in *Ruegeria* sp. TM1040 motility and biofilm formation are controlled by a QS-like system based on a diffusible signal compound with a molecular mass of about 226 Da. *R. mobilis* strain F1926 may utilize the same or a similar QS system as *Ruegeria* sp. TM1040, and since in this case QS and c-di-GMP signalling control the same phenotypes, the two regulation systems are likely connected at some level. In a review of connections between QS and c-di-GMP signalling, Srivastava and Waters (2012) propose that QS signals are generally integrated into the epistatic c-di-GMP signalling system, allowing information about local cell density to be merged with other environmental cues for making a decision between attached and planktonic life. Thus, studying the connection between QS and c-di-GMP signalling may provide further insight into how motility and biofilm formation are controlled.

Most studies of c-di-GMP signalling have approached the role of the compound using bioinformatic and transcriptional tools or genetic manipulation. Few studies actually measured concentrations of c-di-GMP to substantiate their findings (e.g., Weinhouse *et al*., 1997; Waters *et al*., 2008; Merritt *et al*., 2010; Spangler *et al*., 2010). However, due to the ionic nature of c-di-GMP, it used to be difficult to obtain reproducible retention times, as well as sharp symmetrical peaks using conventional

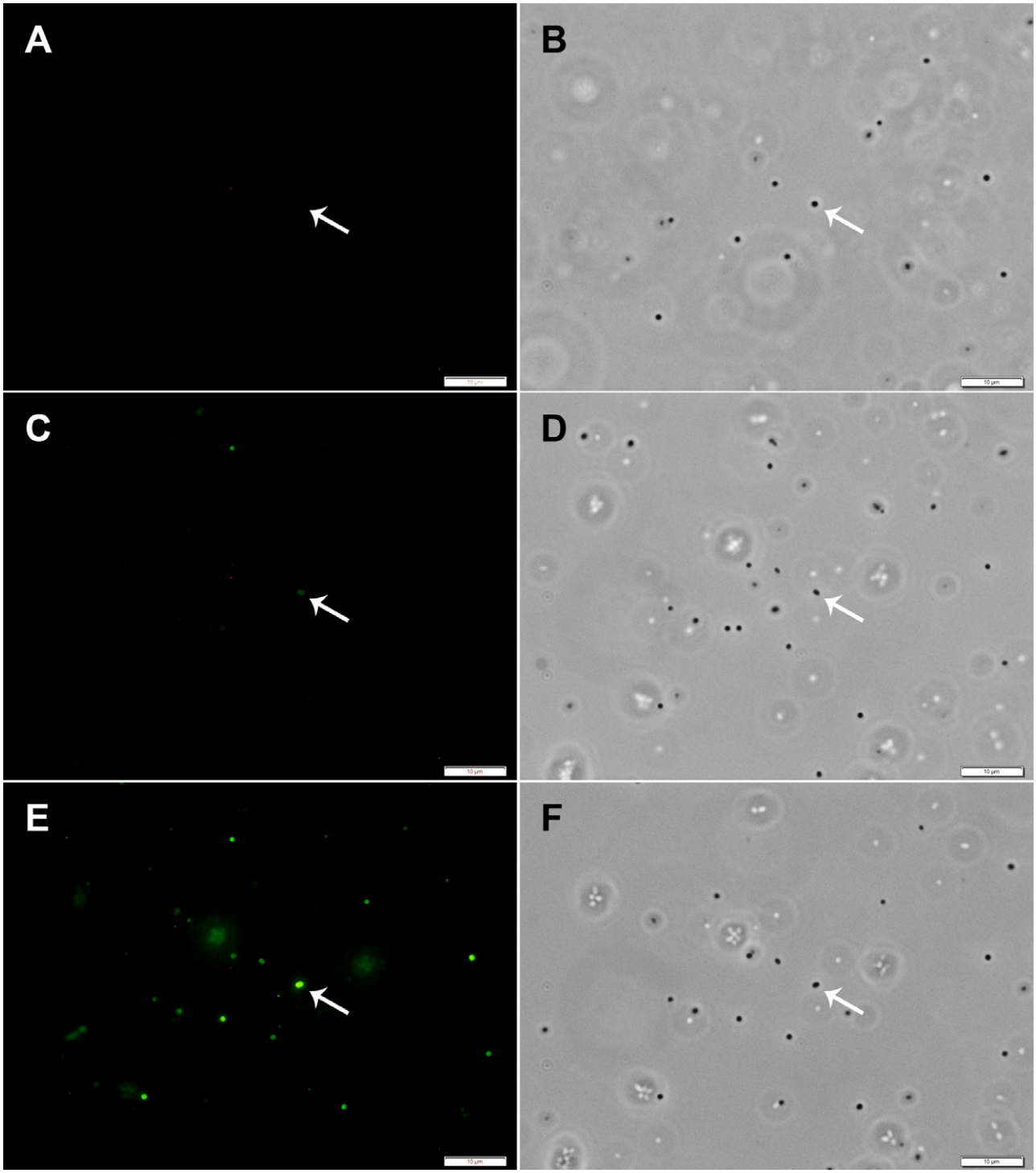


Fig. 6. Time series of *tda*C expression in newly attached *Ruegeria mobilis* F1926 pPDA11. The images were recorded 1 min (A, B), 10 min (B, C) and 20 min (E, F) after preparing the specimen. Fluorescence (A, C, E) and phase contrast (B, D, F) micrographs of the same area were each recorded using the same settings. The focal plane was set right beneath the cover slip to record attached cells. White arrows indicate the position of the same attached cell in all images. Right after preparing the specimen, the freshly attached cells were not fluorescent (A), yet green fluorescence indicating expression of *tda*C was observed in attaching cells after 10 min. All scale bars indicate 10 μm.

reversed phase chromatography (Werner, 1991; Simm *et al*., 2004). LC-MS/MS with tributylamine as ion pair reagent was chosen due to its sensitivity and higher volatility of the tributylamine (Magdenoska *et al*., 2013) when compared with tetrabutylammonium (Witters *et al*., 1997) which reduces background and ion-source pollution (Holcapek *et al*., 2004).

A relative of the *Roseobacter* clade within the α-Proteobacteria, *Caulobacter crescentus*, has a very sophisticated variant of c-di-GMP-mediated 'swim-and-stick' life (Viollier *et al*., 2002; Aldridge *et al*., 2003; Paul *et al*., 2004; Huitema *et al*., 2006; Duerig *et al*., 2009; Abel *et al*., 2011). The cell cycle comprises a stage of flagellated swarmer cells, in which replication is inhibited via low c-di-GMP levels, and a stage of sessile, stalked cells that form new swarmer cells at their non-attached end. In the stalked cells, c-di-GMP is unequally distributed as an effect of antipodal location of GGDEF and EAL domain proteins, restricting cell division to the non-attached pole. In the present study, star-shaped rosettes were not formed in a mutant with lowered c-di-GMP levels, indicating an involvement of the compound in producing that phenotype. Rosette formation in *Roseobacter* clade species could be the result of a process involving polar differences in c-di-GMP contents, where cell division is possibly restricted to one pole of the rosette-forming cell and its daughter cells.

Bruhn and colleagues (2007) demonstrated that cells of *Phaeobacter* sp. 27-4 from static cultures attached better to a glass surface than cells from shaken cultures. The same pattern was observed in this study. However, comparing attachment between shaken cultures of wild type and mutants with altered c-di-GMP levels, we found that initial attachment was compromised both by increased and decreased c-di-GMP levels. Miller and Belas (2006) demonstrated that in *Ruegeria* sp. TM1040, motility is crucial for initiating the *Ruegeria*–dinoflagellate symbiosis. Consequently, the strain with increased c-di-GMP levels may have a reduced capability of attaching to substrates because motility was repressed. Interestingly, attachment of the pYhjH-carrying strain with decreased c-di-GMP and increased motility was reduced as well, even if not to the same extent. This might be caused by a reduced ability to turn off flagellar motility in response to initial surface contact by c-di-GMP-mediated allosteric inhibition of the flagellar motor as known from *C. crescentus* (Christen *et al*., 2007).

Antibacterial activity, production of the brown pigment and TDA production were reduced by decreased intracellular c-di-GMP levels, suggesting that the association between the biofilm phenotype and TDA production could be c-di-GMP-mediated. This led us to study expression of *tda*C as indicator of TDA production on single-cell level, and we found that *tda*C is expressed differently within different cells of the same cultures. This argues against a regulation mechanism based exclusively on community level. Two regulation mechanisms for TDA production were identified on community level. QS was found to activate production of TDA in *P. gallaeciensis* (Berger *et al*., 2011), and also TDA itself was observed to act as an autoinducer, causing increased expression of genes necessary for its own production (Geng and Belas, 2010a). However, both mechanisms fail to explain how *tda*C can be expressed differently in adjacent cells. The hypothesis that TDA production is regulated by c-di-GMP provides an alternative explanation for how TDA production can be different in cells within the same culture, and for how it can, on single-cell level, be spontaneously induced by attachment despite absence of TDA.

The *tda*C gene is expressed only in biofilms or aggregates, and its expression could be initiated in planktonic cells by physical attachment to a surface. Thus, *tda*C was expressed where high or rising levels of c-di-GMP would probably be found. Geng and Belas (2011) showed that the TdaA protein, a LysR-type transcriptional regulator, binds to the *tda*C promoter and activates *tda*C expression. The ligands of LysR-type transcriptional regulator are small molecules (Schell, 1993), and it could be speculated whether c-di-GMP is the ligand of TdaA, or whether intermediate steps are involved.

In conclusion, our study adds organisms from the *Roseobacter* clade to the list of bacteria that use c-di-GMP as a key secondary messenger. Notably, c-di-GMP may be the key molecule in the often described 'stick-and-swim' lifestyle of several roseobacters.

Experimental procedures

Strains, plasmids and media

An overview of strains and plasmids is provided in Table 3. *R. mobilis* F1926 was isolated from the central Indian Ocean (coordinates −31.4061, 91.17758) during the Galathea III expedition and was identified by its 16S rRNA gene sequence (L. Gram, P. D'Alvise, C. Porsby, J. Melchiorsen, J. Heilmann, M. Jensen, *et al*., unpublished data) using procedures described in Gram and colleagues (2010). The strain was revived from frozen stock cultures (−80°C) on half-strength marine agar [½MA; 27.6 g Difco 212185 marine agar (Difco Laboratories, Detroit, MI, USA), 15 g sea salts (Instant Ocean, Vernon, Canada), 7.5 g agar, 1 l deionized water]. Plasmids pYedQ and pYhjH were obtained from Tim Tolker-Nielsen (University of Copenhagen) and electroporated into *R. mobilis* F1926, as described below. Half-strength yeast-tryptone-sea-salts broth (½YTSS) [2 g yeast extract (Bacto Laboratories, Sydney, Australia), 1.25 g tryptone (Bacto Laboratories, Sydney, Australia), 20 g sea salts (Sigma-Aldrich, St. Louis, MO, USA), 1l deionized water] and agar (Gonzalez *et al*., 1996) containing 50 μg ml^{-1} of tetracycline were used for selecting transconjugants after the electroporations and for routine culturing of the plasmid-carrying

Table 3. Bacterial strains and plasmids used in the present study.

Strain or plasmid	Genotype or relevant markers	Source or reference
Strain		
Ruegeria mobilis F1926	Wild type	Gram *et al.* 2010; L. Gram, P. D'Alvise, C. Porsby, J. Melchiorsen, J. Heilmann, M. Jensen, et al., unpubl. data
E. coli TransforMax EC100D *pir*$^{+}$	*pir*$^{+}$	Epicentre
Plasmids		
pRK415	Conjugative broad host range vector, Tet^{R}	Keen *et al.* 1988
pPDA11	tdaCp::*gfp* ligated into pRK415, Tet^{R}	This study; D'Alvise *et al* 2012
pYedQ	*E. coli* gene *yed*Q (diguanylyl cyclase) ligated into pRK404A, Tet^{R}	Ausmees *et al.* 2001; Gjermansen *et al.* 2006
pYhjH	*E. coli* gene *yhj*H (c-di-GMP-specific phosphodiesterase) ligated into pBBR1MCS-3, Tet^{R}	Gjermansen *et al.* 2006
pRK404A	Standard broad host range cloning vector, Tet^{R}, control for pYedQ	Ditta *et al.* 1985
pBBR1MCS-3	Standard broad host range cloning vector, Tet^{R}, control for pYhjH	Kovach *et al.* 1995

mutants. Cultures for microscopy and for chemical measurements of c-di-GMP were grown in full strength MB that contained 50 μg ml^{-1} of tetracycline for the plasmid-carrying mutants. For inhibition testing and TDA analysis, all strains were cultured in MB without addition for 3 days. A 1.5 YTSS with and without tetracycline was used for biofilm and attachment assays. All cultures were grown as 20 ml batches in 250 ml glass bottles at 25°C, except as noted otherwise, and shaking velocity was 200 r.p.m.

Genome sequencing

Genomic DNA was obtained from strain F1926 by successive phenol-chloroform purification steps (Sambrook and Russel, 2001). Mate pair library preparation and Illumina Hi Seq 2000 (Illumina, San Diego, CA, USA) sequencing were carried out by the Bejing Genomic Institute (Shenzhen, China). Contigs were assembled using CLC Genomic Workbench (CLC Bio, Aarhus, Denmark). The genomic DNA sequence has been submitted to the National Center for Biotechnology Information (NCBI) database under accession number AQCH00000000.1.

Detection of genes with GGDEF and EAL domains

The genome draft was annotated by NCBI using the PGAAP, and genes encoding diguanylate cyclases and phosphodiesterases were identified (Table 1). The sequences of these genes were used to search the Pfam database (Punta *et al.*, 2012), and the e-values of the Pfam-identification of the GGDEF- and EAL-domain proteins are stated in Table 1.

Electroporations

The electroporation method was adapted from Miller and Belas (2006). Recipient cells were grown in 50 ml ½YTSS (*R. mobilis*) or LB medium (*E. coli*, 37°C) until OD600 was about 0.5, chilled on ice for 30 min, harvested by centrifugation at 2380 × *g*, washed twice in 10 ml autoclaved, ice-cold MilliQ-water (Merck Millipore, Billerica, MA, USA) and resuspended in 0.5 ml ice-cold 10% glycerol. Aliquots of 70 μl were stored at –80°C until use. Electrocompetent cells were mixed with 180–230 ng plasmid DNA, incubated 30 min on ice, transferred to a 0.2 cm electroporation cuvette (165–2086 Biorad, Hercules, CA, USA) and electroporated at 2.5 kV cm^{-1}, 200 Ω, 25 μF using a Biorad Gene Pulser (Biorad). The cells were immediately transferred to liquid growth medium without antibiotics, recovered for 2–4 h and plated on selective agar.

c-di-GMP extraction

Shaken and static cultures of *R. mobilis* F1926 wild type and the plasmid-carrying mutants were inoculated from OD-adjusted overnight precultures and grown for 24 h in triplicates. The cultures were cooled on ice, and static cultures were shaken briefly to break up the biofilms. One millilitre was sampled and vortexed vigorously to further break up aggregates before measuring OD600. Cultures were harvested (5000 × *g*), and the pellets were extracted with 10 ml 75% (v/v) boiling ethanol/water containing 10 μM HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid]. The pellets were resuspended in 75% ethanol, and the suspensions were left in an 80°C water bath for 5 min. Subsequently, the suspensions were centrifuged (4250 × g), and the supernatants were evaporated to dryness under nitrogen. The samples were dissolved in 100 μl mobile phase A and filtered through 0.2 μm PTFE (polytetrafluoroethylene) hydrophilic filters before analysis.

Ion pair UHPLC-MS/MS analysis of c-di-GMP

The analysis was carried out on an Agilent 1290 (Agilent, Torrance, CA, USA) binary UHPLC system coupled with a 6460 triple quadrupole mass spectrometer (Agilent). The MS was operated in negative electrospray using the $[M\text{-}H]^{-}$ m/z 689.1 as parent ion, and m/z 149.9 and 537.9 as quantifier and qualifier ions respectively. Separation of 10 μL samples was performed by ion pair chromatography, as described

in detail in (Magdenoska *et al*., 2013), using 10 mM tributylamine as ion pair reagent. The gradient used was 0–12 min 0-50% B, 12–12.5 min 50-100% B, 12.5–13 min 100% B, 13–13.1 min 100–0% B, 13.1–18 min 0% B. Three hundred and sixty milliliters of shaken cultures of the pYhjH-carrying mutant were split into eighteen 50 ml falcon tube and used for matrix-matched calibration. One mg/ml c-di-GMP in water was used to prepare the spiking solutions. After centrifugation and removal of the supernatants, 75% boling EtOH was added to the tubes followed by spiking with, 100 μl of 0 ng/ml, 35 ng/ml, 100 ng/ml, 200 ng/ml, 600 ng/ml and 1000 ng/ml c-di-GMP standard in triplicates. The spiked cultures were extracted and prepared for analysis as described above. The amount of c-di-GMP detected in the blank was subtracted from the spiked calibrants, and the analysis was calibrated by linear regression ($r^2 = 0.993$). To obtain a relative estimate of c-di-GMP concentrations per cellular biomass, c-di-GMP concentrations were divided by the measured OD600 of the original cultures.

Biofilm and attachment assay

Biofilm formation in *R. mobilis* F1926 wild type, F1926 pYedQ and F1926 pYhjH was measured by a crystal violet method (O'Toole *et al*., 2000). Briefly, shaken precultures were diluted with fresh medium to an OD600 of 0.1, pipetted into a 96-well microtiter plate and incubated for 24 h. Culture liquid was removed, and biofilms were washed and stained in 1% (w/v) crystal violet solution. After washing, the crystal violet was extracted from the stained biofilms with 96% ethanol and quantified by measuring absorption at 590 nm.

Attachment to an inert surface was measured in a modified crystal violet assay. Static cultures were grown in 96-well plates as described above. Shaken cultures were grown in glass bottles for 24 h; OD600 was adjusted to 1.0, and 200 μl were pipetted into the wells of a micotiter plate. A lid with 96 polystyrene pegs (Innovotech, Edmonton, Canada) was placed on the plate, and the cells were allowed to attach to the pegs for 1 min. Adherent biofilms on the pegs were washed twice by dipping into water, dried for 5 min and stained in crystal violet solution. After triple washing in water, the crystal violet was extracted from the stained biofilms on the pegs in each 200 μl of ethanol, and absorption was measured at 590 nm. Both assays were conducted in duplicates and reproduced independently.

Measurement of antibacterial activity

Inhibition of *Vibrio anguillarum* 90-11-287 in a standard well-diffusion assay was measured as an approximation to TDA production as adapted from (Hjelm *et al*., 2004). *V. anguillarum* 90-11-287 was grown in MB for 1 day at 25°C with aeration at 200 r.p.m. A 50 μl of the *V. anguillarum* preculture was added to 50 ml molten Instant Ocean agar [1.5 g of Instant Ocean sea salts, 0.1 g of casamino acids (Bacto Laboratories, Sydney, Australia), 0.2 g glucose, 0.5 of g agar] at 41.5°C and poured into a 14 cm Petri dish. Wells of 6 mm diameter were punched into the solidified agar and filled with 50 μl of *R. mobilis* F1926 culture supernatant. The assay was incubated for 1 day at 25°C, and diameters of inhibition zones were measured. The results are based on two independent replicates.

TDA extraction and analysis

A 1 ml of each MB culture was mixed with 3 ml ethyl acetate containing 1% formic acid and extracted for 30 min on a shaking table at room temperature. A 2.5 ml of the organic phase was evaporated to dryness under nitrogen flow at 35°C and redissolved in 100 μl 85% acetonitrile/15% MilliQ water. Blank medium samples spiked with 0.13–100 μM pure TDA (BioViotica, Dransfeld, Germany) served as standard. UHPLC-TOFMS analysis was conducted on an Agilent 1290 UHPLC coupled to an Agilent 6550 qTOF (Agilent) equipped with a dual electrospray source. Separation was performed at 40°C on a 2.1 mm ID, 50 mm, 1.8 μm of Eclipse Plus C_{18} (Agilent) column using a water-acetonitrile gradient solvent system, with both water and acetonitrile containing 20 mM of formic acid. Using a flow of 0.8 ml/min, the gradient was started at 15% acetonitrile and increased to 60% acetonitrile within 1.8 min, then to 100% in 0.2 min, keeping this for 0.8 min, returning to 15% acetonitrile in 0.2 min and equilibrating for the next sample in 1.5 min (total runtime is 4.5 min). TDA was determined in ESI^+ mode and quantified from its $[M + H]^+$ ion 212.9674 ± 0.005 with the same retentions as the authentic standard (0.97 min). Quantification was done using 1/x weighted linear regression based on the peak area in the MassHunter Quant 5.0 software (Agilent). Duplicate cultures were used for TDA analysis, and the experiment was reproduced independently.

Statistics

Cellular c-di-GMP concentrations and TDA concentrations were compared by *t*-tests. Differences in average crystal violet absorption values in the biofilm and attachment assays were examined by one-way ANOVA with Tukey's multiple comparison test using the software Prism version 4.03 (GraphPad Software, La Jolla, CA, USA). All average values and standard deviations are based on biological replicates.

Construction of a tda*Cp::*gfp *reporter fusion*

A transcriptional fusion between the promoter of *tda*C and a promoterless *gfp* gene was constructed similarly to pHG1011 (Geng and Belas, 2010a). The promoter sequence of the *tda*C gene was amplified using the primers ptdacF (5′-GTCCCAGAGACCAACGCAATGAGTAAAGGAGAAGAA-3′) and ptdacR (5′-TTCTTCTCCTTTACTCATTGCGTTGGTCTC TGGGAC-3′). The *gfp* open reading frame in pAKN137 (Lambertsen *et al*., 2004) was amplified using the primers gfpF (5′-GTCCCAGAGACCAACGCAATGAGTAAAGGAGAA GAA-3′) and gfpR (5′-TGATAAGCTTTTATTTGTATAGTTCAT CCATGCCATGT-3′). Primer ptdacF created a PstI-restriction site, and gfpR created a HindIII-site. Primer ptdacR and gfpF created identical 36-bp-sequences in the adjacent ends of the two amplicons, each containing the end of the promoter and the start of the *gfp* open reading frame. This allowed seamless cloning of promoter and *gfp* gene by overlap-extension polymerase chain reaction. The product was cloned into the broad-host range vector pRK415 (Keen *et al*., 1988) after both had been digested with PstI and HindIII (New England Biolabs, Ipswich, MA, USA) to yield plasmid pPDA11.

Microscopy

Shaken and static cultures of *R. mobilis* F1926 wild type and the mutants carrying pYedQ and pYhjH were grown in duplicates for 24 h and were compared by phase contrast microscopy. Before specimen preparation, static cultures were agitated briefly to break the biofilms into smaller pieces for sampling. Images that were representative of the specimen were recorded. Screening the whole specimen, rosette formation, and the proportion of single cells and cells in biofilms were registered. The proportion of motile cells was estimated visually. Absence of motile cells or rosettes was stated if no motile cell or rosette was observed in either of the duplicate samples.

Gfp-fluorescence of *R. mobilis* F1926 pPDA11 was detected by microscopy using a long-pass fluorescence cube (Olympus, Tokyo, Japan; WIB ex. 460–490, em. > 515). Shaken and static cultures were grown for 3 days. Again, static cultures were briefly shaken before specimen preparation. Representative fluorescence micrographs were recorded with 1.5 s exposure, and a phase contrast image of the same area was recorded right thereafter. To record a time series showing the onset of *gfp* expression in response to attachment, a 3 day old shaken culture was diluted 1:2 with fresh medium and grown for 4 h at the same conditions. A specimen was prepared, and a time series of fluorescence micrographs was recorded as described above. For better display, contrast of the phase, contrast pictures and brightness of the fluorescence micrographs were enhanced using Photoshop (Adobe, San Jose, CA, USA). The same adjustments were made on all images of the same type.

Motility agar

Motility agar was prepared based on half-strength and 1/10-strength YTSS (0.4 g Bacto Yeast extract, 0.25 g Bacto Tryptone, 20 g Sigma sea salts, 1l deionized water) with different agar percentages (0.5, 0.45, 0.4, 0.35, 0.3 and 0.2%) and 5 ml/l 1% tetrazolium red solution, and Stab agar tubes and Petri dishes were prepared. The Petri dishes were inoculated with 5 μl of OD-adjusted (OD600 = 0.5) 1.5 YTSS precultures, and the stab agar cultures were inoculated with a needle dipped into the same OD-adjusted precultures. The diameters of growth were compared after 1, 2 and 5 days.

Acknowledgements

We thank Tim Tolker-Nielsen (University of Copenhagen) for providing plasmids pYedQ and pYhjH, as well as the vector controls, and we thank Jens Bo Andersen (University of Copenhagen) for technical advice with the construction of plasmid pPDA11. We are grateful to Agilent technologies for the Thought Leader Donation of the UHPLC-qTOF system.

This work was funded by the Danish Research Council for Technology and Production Sciences (project 09-066524).

References

Abel, S., Chien, P., Wassmann, P., Schirmer, T., Kaever, V., Laub, M.T., *et al*. (2011) Regulatory cohesion of cell cycle and cell differentiation through inter linked phosphorylation and second messenger networks. *Mol Cell* **43:** 550–560.

Alavi, M., Miller, T., Erlandson, K., Schneider, R., and Belas, R. (2001) Bacterial community associated with *Pfiesteria*-like dinoflagellate cultures. *Env Microbiol* **3:** 380–396.

Aldridge, P., Paul, R., Goymer, P., Rainey, P., and Jenal, U. (2003) Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus*. *Mol Microbiol* **47:** 1695–1708.

Allgaier, M., Uphoff, H., Felske, A., and Wagner-Dobler, I. (2003) Aerobic anoxygenic photosynthesis in *Roseobacter* clade bacteria from diverse marine habitats. *Appl Environ Microbiol* **69:** 5051–5059.

Ausmees, N., Mayer, R., Weinhouse, H., Volman, G., Amikam, D., Benziman, M., and Lindberg, M. (2001) Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity. *FEMS Microbiol Let* **204:** 163–167.

Belas, R., Horikawa, E., Aizawa, S.I., and Suvanasuthi, R. (2009) Genetic determinants of *Silicibacter* sp. TM1040 motility. *J Bacteriol* **191:** 4502–4512.

Berger, M., Neumann, A., Schulz, S., Simon, M., and Brinkhoff, T. (2011) Tropodithietic acid production in *Phaeobacter gallaeciensis* is regulated by N-acyl homoserine lactone-mediated quorum sensing. *J Bacteriol* **193:** 6576–6585.

Brinkhoff, T., Bach, G., Heidorn, T., Liang, L.F., Schlingloff, A., and Simon, M. (2004) Antibiotic production by a *Roseobacter* clade-affiliated species from the German Wadden Sea and its antagonistic effects on indigenous isolates. *Appl Environ Microbiol* **70:** 2560–2565.

Brinkhoff, T., Giebel, H.A., and Simon, M. (2008) Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch Microbiol* **189:** 531–539.

Bruhn, J.B., Nielsen, K.F., Hjelm, M., Hansen, M., Bresciani, J., Schulz, S., and Gram, L. (2005) Ecology, inhibitory activity, and morphogenesis of a marine antagonistic bacterium belonging to the *Roseobacter* clade. *Appl Environ Microbiol* **71:** 7263–7270.

Bruhn, J.B., Haagensen, J.A.J., Bagge-Ravn, D., and Gram, L. (2006) Culture conditions of *Roseobacter* strain 27-4 affect its attachment and biofilm formation as quantified by real-time PCR. *Appl Environ Microbiol* **72:** 3011–3015.

Bruhn, J.B., Gram, L., and Belas, R. (2007) Production of antibacterial compounds and biofilm formation by *Roseobacter* species are influenced by culture conditions. *Appl Environ Microbiol* **73:** 442–450.

Buchan, A., Neidle, E.L., and Moran, M.A. (2001) Diversity of the ring-cleaving dioxygenase gene *pca*H in a salt marsh bacterial community. *Appl Environ Microbiol* **67:** 5801–5809.

Buchan, A., Gonzalez, J.M., and Moran, M.A. (2005) Overview of the marine *Roseobacter* lineage. *Appl Environ Microbiol* **71:** 5665–5677.

Charlson, R.J., Lovelock, J.E., Andreae, M.O., and Warren, S.G. (1987) Oceanic phytoplankton, atmospheric sulfur, cloud albedo and climate. *Nature* **326:** 655–661.

Christen, M., Christen, B., Folcher, M., Schauerte, A., and Jenal, U. (2005) Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *J Biol Chem* **280:** 30829–30837.

Christen, M., Christen, B., Allan, M.G., Folcher, M., Jeno, P., Grzesiek, S., and Jenal, U. (2007) DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*. *Proc Natl Acad Sci USA* **104:** 4112–4117.

Cotter, P.A., and Stibitz, S. (2007) c-di-GMP-mediated regulation of virulence and biofilm formation. *Curr Opin Microbiol* **10:** 17–23.

D'Alvise, P.W., Melchiorsen, J., Porsby, C.H., Nielsen, K.F., and Gram, L. (2010) Inactivation of *Vibrio anguillarum* by attached and planktonic *Roseobacter* cells. *Appl Environ Microbiol* **76:** 2366–2370.

D'Alvise, P.W., Lillebo, S., Prol-Garcia, M.J., Wergeland, H.I., Nielsen, K.F., Bergh, O., and Gram, L. (2012) *Phaeobacter gallaeciensis* reduces *Vibrio anguillarum* in cultures of microalgae and rotifers, and prevents vibriosis in cod larvae. *PLoS ONE* **7:** e43996.

Ditta, G., Schmidhauser, T., Yakobson, E., Lu, P., Liang, X.W., Finlay, D.R., *et al*. (1985) Plasmids related to the broad host range vector pRK290, useful for gene cloning and for monitoring gene expression. *Plasmid* **13:** 149–153.

Duerig, A., Abel, S., Folcher, M., Nicollier, M., Schwede, T., Amiot, N., *et al*. (2009) Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. *Genes Dev* **23:** 93–104.

Geng, H.F., and Belas, R. (2010a) Expression of tropodithietic acid biosynthesis is controlled by a novel autoinducer. *J Bacteriol* **192:** 4377–4387.

Geng, H.F., and Belas, R. (2010b) Molecular mechanisms underlying roseobacter-phytoplankton symbioses. *Curr Opin Biotech* **21:** 332–338.

Geng, H.F., and Belas, R. (2011) TdaA regulates tropodithietic acid synthesis by binding to the *tda*C promoter region. *J Bacteriol* **193:** 4002–4005.

Geng, H.F., Bruhn, J.B., Nielsen, K.F., Gram, L., and Belas, R. (2008) Genetic dissection of tropodithietic acid biosynthesis by marine roseobacters. *Appl Environ Microbiol* **74:** 1535–1545.

Gjermansen, M., Ragas, P., and Tolker-Nielsen, T. (2006) Proteins with GGDEF and EAL domains regulate *Pseudomonas putida* biofilm formation and dispersal. *FEMS Microbiol Lett* **265:** 215–224.

Gonzalez, J.M., and Moran, M.A. (1997) Numerical dominance of a group of marine bacteria in the alpha-subclass of the class *Proteobacteria* in coastal seawater. *Appl Environ Microbiol* **63:** 4237–4242.

Gonzalez, J.M., Whitman, W.B., Hodson, R.E., and Moran, M.A. (1996) Identifying numerically abundant culturable bacteria from complex communities: an example from a lignin enrichment culture. *Appl Environ Microbiol* **62:** 4433–4440.

Gonzalez, J.M., Simo, R., Massana, R., Covert, J.S., Casamayor, E.O., Pedros-Alio, C., and Moran, M.A. (2000) Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl Environ Microbiol* **66:** 4237–4246.

Gram, L., Melchiorsen, J., and Bruhn, J.B. (2010) Antibacterial activity of marine culturable bacteria collected from a global sampling of ocean surface waters and surface swabs of marine organisms. *Mar Biotechnol* **12:** 439–451.

Hengge, R. (2009) Principles of c-di-GMP signalling in bacteria. *Nat Rev Microbiol* **7:** 263–273.

Hjelm, M., Bergh, O., Riaza, A., Nielsen, J., Melchiorsen, J., Jensen, S., *et al*. (2004) Selection and identification of autochthonous potential probiotic bacteria from turbot larvae (*Scophthalmus maximus*) rearing units. *Syst Appl Microbiol* **27:** 360–371.

Holcapek, M., Volna, K., Jandera, P., Kolarova, L., Lemr, K., Exner, M., and Cirkva, A. (2004) Effects of ion-pairing reagents on the electrospray signal suppression of sulphonated dyes and intermediates. *J Mass Spectrom* **39:** 43–50.

Huitema, E., Pritchard, S., Matteson, D., Radhakrishnan, S.K., and Viollier, P.H. (2006) Bacterial birth scar proteins mark future flagellum assembly site. *Cell* **124:** 1025–1037.

Keen, N.T., Tamaki, S., Kobayashi, D., and Trollinger, D. (1988) Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. *Gene* **70:** 191–197.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., and Peterson, K.M. (1995) 4 new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance casettes. *Gene* **166:** 175–176.

Lai, Q.L., Yuan, J., Li, F.Y., Zheng, T.L., and Shao, Z.Z. (2010) *Ruegeria pelagia* is a later heterotypic synonym of *Ruegeria mobilis*. *Int J Syst Evol Microbiol* **60:** 1918–1920.

Lambertsen, L., Sternberg, C., and Molin, S. (2004) Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins. *Environ Microbiol* **6:** 726–732.

McDougald, D., Rice, S.A., Barraud, N., Steinberg, P.D., and Kjelleberg, S. (2012) Should we stay or should we go: mechanisms and ecological consequences for biofilm dispersal. *Nat Rev Microbiol* **10:** 39–50.

Magdenoska, O., Martinussen, J., Thykaer, J., and Nielsen, K.F. (2013) Dispersive solid phase extraction combined with ion-pair ultra high-performance liquid chromatography tandem mass spectrometry for quantification of nucleotides in *Lactococcus lactis*. *Anal Biochem* **440:** 166–177.

Matilla, M.A., Travieso, M.L., Ramos, J.L., and Ramos-Gonzalez, M.I. (2011) Cyclic diguanylate turnover mediated by the sole GGDEF/EAL response regulator in Pseudomonas putida: its role in the rhizosphere and an analysis of its target processes. *Env Microbiol* **13:** 1745–1766.

Merritt, J.H., Ha, D.G., Cowles, K.N., Lu, W.Y., Morales, D.K., Rabinowitz, J., *et al*. (2010) Specific control of *Pseudomonas aeruginosa* surface-associated behaviors by two c-di-GMP diguanylate cyclases. *mBio* **1:** 00183–10.

Miller, T.R., and Belas, R. (2006) Motility is involved in *Silicibacter* sp. TM1040 interaction with dinoflagellates. *Environ Microbiol* **8:** 1648–1659.

Moran, M.A., and Hodson, R.E. (1994) Dissolved humic substances of vascular-plant origin in a coastal marine environment. *Limnol Oceanogr* **39:** 762–771.

Moran, M.A., Gonzalez, J.M., and Kiene, R.P. (2003) Linking a bacterial taxon to sulfur cycling in the sea: studies of the marine *Roseobacter* group. *Geomicrobiol J* **20:** 375–388.

Moran, M.A., Buchan, A., Gonzalez, J.M., Heidelberg, J.F., Whitman, W.B., Kiene, R.P., *et al*. (2004) Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* **432:** 910–913.

Moran, M.A., Belas, R., Schell, M.A., Gonzalez, J.M., Sun, F., Sun, S., *et al*. (2007) Ecological genomics of marine roseobacters. *Appl Environ Microbiol* **73:** 4559–4569.

Newell, P.D., Yoshioka, S., Hvorecny, K.L., Monds, R.D., and O'Toole, G.A. (2011) Systematic analysis of diguanylate cyclases that promote biofilm formation by Pseudomonas fluorescens Pf0-1. *J Bacteriol* **193:** 4685–4698.

Newton, R.J., Griffin, L.E., Bowles, K.M., Meile, C., Gifford, S., Givens, C.E., *et al*. (2010) Genome characteristics of a generalist marine bacterial lineage. *ISME J* **4:** 784–798.

O'Toole, G., Kaplan, H.B., and Kolter, R. (2000) Biofilm formation as microbial development. *Ann Rev Microbiol* **54:** 49–79.

Paul, R., Weiser, S., Amiot, N.C., Chan, C., Schirmer, T., Giese, B., and Jenal, U. (2004) Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes Dev* **18:** 715–727.

Planas, M., Perez-Lorenzo, M., Hjelm, M., Gram, L., Fiksdal, I.U., Bergh, O., and Pintado, J. (2006) Probiotic effect *in vivo* of *Roseobacter* strain 27-4 against *Vibrio* (*Listonella*) *anguillarum* infections in turbot (*Scophthalmus maximus* L.) larvae. *Aquacult* **255:** 323–333.

Porsby, C.H., Nielsen, K.F., and Gram, L. (2008) *Phaeobacter* and *Ruegeria* species of the *Roseobacter* clade colonize separate niches in a Danish turbot (*Scophthalmus maximus*)-rearing farm and antagonize *Vibrio anguillarum* under different growth conditions. *Appl Environ Microbiol* **74:** 7356–7364.

Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., *et al*. (2012) The Pfam protein families database. *Nucleic Acids Res* **40:** D290–D301.

Ruiz-Ponte, C., Cilia, V., Lambert, C., and Nicolas, J.L. (1998) *Roseobacter gallaeciensis* sp. nov., a new marine bacterium isolated from rearings and collectors of the scallop *Pecten maximus*. *Int J System Bacteriol* **48:** 537–542.

Ryjenkov, D.A., Tarutina, M., Moskvin, O.V., and Gomelsky, M. (2005) Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol* **187:** 1792–1798.

Sambrook, J., and Russel, D. (2001) *Molecular Cloning – A Laboratory Manual*. 3rd edn. Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press.

Schell, M.A. (1993) Molecular biology of the LysR family of transcriptional regulators. *Ann Rev Microbiol* **47:** 597–626.

Schmidt, A.J., Ryjenkov, D.A., and Gomelsky, M. (2005) The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J Bacteriol* **187:** 4774–4781.

Shiba, T., Simidu, U., and Taga, N. (1979) Another aerobic bacterium which contains bacteriochlorophyll-A. *B Jpn Soc Sci Fish* **45:** 801.

Simm, R., Morr, M., Kader, A., Nimtz, M., and Romling, U. (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* **53:** 1123–1134.

Slightom, R.N., and Buchan, A. (2009) Surface colonization by marine Roseobacters: integrating genotype and phenotype. *Appl Environ Microbiol* **75:** 6027–6037.

Sorokin, D.Y., and Lysenko, A.M. (1993) Heterotrophic bacteria from the Black Sea oxidizing reduced sulfur compounds to sulfate. *Microbiology* **62:** 594–602.

Spangler, C., Bohm, A., Jenal, U., Seifert, R., and Kaever, V. (2010) A liquid chromatography-coupled tandem mass spectrometry method for quantitation of cyclic di-guanosine monophosphate. *J Microbiol Methods* **81:** 226–231.

Srivastava, D., and Waters, C.M. (2012) A tangled web: regulatory connections between quorum sensing and cyclic Di-GMP. *J Bacteriol* **194:** 4485–4493.

Sule, P., and Belas, R. (2012) A novel inducer of *Roseobacter* motility is also a disruptor of algal symbiosis. *J Bacteriol* **195:** 637–646.

Tamayo, R., Pratt, J.T., and Camilli, A. (2007) Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Ann Rev Microbiol* **61:** 131–148.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28:** 2731–2739.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) Clustal-W – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22:** 4673–4680.

Viollier, P.H., Sternheim, N., and Shapiro, L. (2002) Identification of a localization factor for the polar positioning of bacterial structural and regulatory proteins. *Proc Natl Acad Sci USA* **99:** 13831–13836.

Waters, C.A., Lu, W.Y., Rabinowitz, J.D., and Bassler, B.L. (2008) Quorum sensing controls biofilm formation in *Vibrio cholerae* through modulation of cyclic Di-GMP levels and repression of vpsT. *J Bacteriol* **190:** 2527–2536.

Weinhouse, H., Sapir, S., Amikam, D., Shilo, Y., Volman, G., Ohana, P., and Benziman, M. (1997) c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum*. *FEBS Lett* **416:** 207–211.

Werner, A. (1991) Analysis of nucleotides, nucleosides, nucleobases in cells by ion-pair reversed-phase Hplc. *Chromatographia* **31:** 401–410.

Wietz, M., Gram, L., Jorgensen, B., and Schramm, A. (2010) Latitudinal patterns in the abundance of major marine bacterioplankton groups. *Aquat Microb Ecol* **61:** 179–189.

Witters, E., VanDongen, W., Esmans, E.L., and VanOnckelen, H.A. (1997) Ion-pair liquid chromatography electrospray mass spectrometry for the analysis of cyclic nucleotides. *J Chromatogr B* **694:** 55–63.

Wolfe, A.J., and Visick, K.L. (2008) Get the message out: cyclic-Di-GMP regulates multiple levels of flagellum-based motility. *J Bacteriol* **190:** 463–475.

Yi, H., Lim, Y.W., and Chun, J. (2007) Taxonomic evaluation of the genera *Ruegeria* and *Silicibacter*: a proposal to transfer the genus *Silicibacter* Petursdottir and Kristjansson 1999 to the genus *Ruegeria* Uchino *et al*. 1999. *Int J Syst Evol Microbiol* **57:** 815–819.

Zan, J.D., Cicirelli, E.M., Mohamed, N.M., Sibhatu, H., Kroll, S., Choi, O., *et al*. (2012) A complex LuxR-LuxI type quorum sensing network in a roseobacterial marine

sponge symbiont activates flagellar motility and inhibits biofilm formation. *Mol Microbiol* **85:** 916–933.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Phylogenetic tree of *Ruegeria mobilis* constructed from 16S rRNA gene sequences. *R. mobilis* F1926 is compared with the *R. mobilis* type strain NBRC 101030 (accession number NR_041454), other *R. mobilis* isolates (accession numbers HQ338144.1, HQ338146.1, HQ338142.1, HQ338140.1, HQ338148.1, HQ338141.1, HQ338145.1, HQ338143.1) and other species from the *Ruegeria/Phaeobacter* subcluster (Newton *et al*., 2010) of the *Roseobacter* clade (accession numbers NR_074151.1, NR_042675.1, NR_043449.1, GU176618.1, HQ_908721.1, NR_029273.1, NR_027609.1, NR_042761.1, AJ536669.1, NR_074150.1, FJ872535.1). *Rhodobacter capsulatus* ATCC11166 and *Rhodobacter sphaeroides* ATTC BAA-808 (accession numbers DQ342320.1 and NR_074174.1) served as outgroup. The 16S sequences were aligned using ClustalW (Thompson *et al*., 1994), and the neighbour-joining tree was constructed using MEGA version 5 (Tamura *et al*., 2011). Numbers at the nodes are bootstrap values from 500 replications.

Fig. S2. Cyclic di-GMP detection in extracts of *Ruegeria mobilis* F1926. Example chromatograms of F1926 pYhjH shaken culture spiked with c-di-GMP (top) and F1926 wild type static culture (bottom). The MRM transition monitored is m/z 689.1→149.9. The counts on the *y*-axis are relative to the highest peak in the respective sample; thus, the figure does not allow for quantitative comparison.

Table S1. Guanylate cyclase and phosphodiesterase genes in TDA-producing and non-TDA-producing *Roseobacter* clade species.

6.3 Paper 3 - Multi-omic profiling of EPO-producing Chinese hamster ovary cell panel reveals metabolic adaptation to heterologous protein production.

Daniel Ley, Ali Kazemi Seresht, Mikael Engmark, **Olivera Magdenoska**, Kristian Fog Nielsen, Helene Faustrup Kildegaard, Mikael Rørdam Andersen,

Paper submitted to Biotechnology and Bioengineering 2015

Multi-omic profiling of EPO-producing Chinese hamster ovary cell panel reveals metabolic adaptation to heterologous protein production

Authors: Daniel Ley[1,2,3,#], Ali Kazemi Seresht[2,#], Mikael Engmark[1,2], Oliviera Magdenoska[1], Kristian Fog Nielsen[1], Helene Faustrup Kildegaard[3], Mikael Rørdam Andersen[1,*].

(1) Department of Systems Biology, Technical University of Denmark, Kgs. Lyngby, Denmark; (2) Cell Culture Technology, Novo Nordisk A/S, Novo Nordisk Park, Måløv, Denmark; (3) The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark.

[#]These authors contributed equally to the work.

*Corresponding author: mr@bio.dtu.dk,

Phone: +45 45 25 26 75, Fax: +45 45 88 41 48

Author contributions: D.L. performed part of the experimental work, developed new analysis methods, analyzed data and wrote the manuscript. A.K.S. performed part of the experimental work, analyzed data and wrote the manuscript. M.E performed part of the experimental work, analyzed data and wrote the manuscript. O.M. developed new methods, performed part of the experimental work and wrote the manuscript. K.F.N. developed new methods and analyzed data. HFK performed part of the experimental work and wrote the manuscript. M.R.A. wrote the manuscript.

ABSTRACT

Chinese hamster ovary (CHO) cells are the preferred production host for many therapeutic proteins. The production of heterologous proteins in CHO cells imposes a burden on the host cell metabolism and impact cellular physiology on a global scale. In this work, a multi-omics approach was applied to study the production of erythropoietin (EPO) in a panel of CHO-K1 cells under growth-limited and unlimited conditions in batch and chemostat cultures. Physiological characterization of the EPO-producing cells included global transcriptome analysis, targeted metabolome analysis, including intracellular pools of glycolytic intermediates, NAD(P)H/NAD(P)$^+$, adenine nucleotide phosphates (ANP) and extracellular concentrations of sugars, organic acids and amino acids. Potential impact of EPO expression on the protein secretory pathway was assessed at multiple stages using quantitative PCR (qPCR), reverse transcription PCR (qRT-PCR), western blots (WB) and global gene expression analysis to assess EPO gene copy numbers, EPO gene expression, intracellular EPO retention and differentially expressed genes functionally related to secretory protein processing, respectively.

We found no evidence supporting the existence of production bottlenecks in energy metabolism (i.e. glycolytic metabolites, NAD(P)H/NAD(P)$^+$ and ANP's) in batch culture or in the secretory protein production pathway (i.e. gene dosage, transcription and post-translational processing of EPO) in chemostat culture at specific productivities up to 5 pg/cell/day. We have shown, that the metabolic response to EPO production includes a redistribution of carbon uptake in batch culture with increased glucose demand, possibly reflecting increased energy requirements from protein production. Furthermore, time-course analysis of high- and low producing clones in chemostat culture revealed rapid adaptation of transcription levels of amino acid catabolic genes in favor of EPO production within 9

generations. Interestingly, the adaptation was followed by an increase in specific EPO productivity.

Keywords: Chinese hamster ovary, erythropoietin, chemostat, metabolomics, transcriptomics, metabolic adaptation.

Introduction

Most biopharmaceutical products like monoclonal antibodies, hormones and blood-related proteins are produced in Chinese hamster ovary (CHO) cells (Walsh 2014). Studies of CHO cells have yielded a basic understanding of mammalian cell biology and driven the development of mammalian cell factories for production of structurally advanced pharmaceutical glycoproteins (Jayapal & Wlaschin 2007). For example, numerous studies have focused on resolving bottlenecks in the protein production and secretory pathway (i.e. transcription, translation, protein translocation, -folding, -modification and -secretion), which limit the cell-specific protein productivity (Kim et al. 2012). Often the production bottleneck is reported to be independent of the heterologous target protein, indicating a general limitation of the secretory protein processing capacity (Jossé et al. 2012), while in some cases the bottleneck is linked to the synthesis of a specific post-translational protein modification (Pybus et al. 2013).

In brief, protein production bottlenecks in CHO cells have been reported at the level of transgene expression (Mason et al. 2012; Jiang et al. 2006; C. J. Lee et al. 2009) and stability of mRNA transcripts (Hung et al. 2010). Numerous studies report a non-linear correlation between mRNA copy numbers and specific protein secretion, indicating limitations of either mRNA translation or post-translational processes (Chusainow et al. 2009; Mead et al. 2009; O'Callaghan et al. 2010; Reisinger et al. 2008; Lattenmayer et al. 2007; Lattenmayer & Loeschel 2007). One study suggests that the translocation of mRNA to

the endoplasmatic reticulum (ER) is limiting protein production in CHO cells (Kang et al. 2014). Other studies have reported bottlenecks in protein folding capacity for specific proteins (Y. Y. Lee et al. 2009; Borth et al. 2005; Hwang et al. 2003; Chung et al. 2004; Mohan & Lee 2010). Furthermore, some studies have found bottlenecks within vesicular transport of proteins from ER to the Golgi apparatus and within exocytotic transport from the trans-Golgi cisternae to the plasma membrane (Peng & Fussenegger 2009; Peng et al. 2011). Finally, for specific glycoproteins, evidence suggest bottlenecks in the processing of N-linked glycan structures (Bolt et al. 2008). In general, all major steps (i.e. transcription, translation, protein translocation, protein folding, protein glycosylation and inter-organelle protein transport) have been argued to be a bottleneck. In many cases, it is possible that the cultivation method is confounding the, as one could expect that different cultivation modes (batch, fed-batch, continuous, various forms of nutrient starvation) will have varying requirements for native protein production, and thus influence the metabolic load on the cell.

Cultivation of recombinant cells is performed in different ways, depending on the goal of the experiment. In batch cultivation, all nutrients are supplied initially in excess, - allowing growth at maximum specific rate with maximum specific nutrient uptake and maximum production of native proteins. In an industrial context, the batch process is of limited use for protein production, since growth and productivity rapidly becomes limited by nutrient availability and by-product inhibition. As an alternative process, where growth, by-product accumulation and nutrient consumption can be controlled, continuous cultures are operated with a constant in-flow of fresh medium, while spent medium, biomass and product is removed at an equal rate. A popular continuous cultivation format for physiological characterization of cells is the chemostat (Bull 2010), which is operated at a constant dilution rate (i.e. rate of medium flow per culture volume), thus ensuring a constant physiochemical environment in the bioreactor. This feature enables the study of effects of single parameters

on the cell physiology. Moreover, the restricted in-flow of fresh medium allows tight control of the growth rate of cultivated cells as availability of nutrients becomes limiting in the culture. The operation at a fixed dilution rate thus enables the normalization of growth rates between parallel cultures, which has been shown to be a prerequisite for global transcriptional profiling as the expression level of many genes is affected by the specific growth rate (Regenberg et al. 2006). Chemostat cultures have been used extensively as a powerful tool for the study of e.g. metabolism, protein production, genetic stability and long-term metabolic adaptation of microorganisms (comprehensively reviewed by Bull, 2010). However, so far only a few studies have described the physiological characterization of CHO cells in chemostat culture (Lee et al. 1998; Nyberg et al. 1999; Hayter et al. 1993; Hayter et al. 1992).

The ‘omics technologies (e.g. genomics, transcriptomics, proteomics, metabolomics, glycomics and fluxomics) provide systems-level data on the intracellular state of a biological system crucial to elucidate the molecular basis of CHO cell physiology (reviewed by Kildegaard et al. 2013). Comparative analysis of ‘omics data gathered under specific physiological conditions has revealed differentially regulated molecular mechanisms responsible for desirable phenotypes in isogenic clone populations and guided the design of improved cell factories (Yee et al. 2009; Smales et al. 2004; Chong et al. 2010; Sengupta et al. 2011).

The metabolic burden imposed by heterologous protein production in mammalian cells is still not well characterized and thus may offer opportunities for further improvement of protein productivity. A recent study by Niklas et al. 2013 comparing human cells expressing α_1-antitrypsin found increased anabolic demand for RNA and lipids in protein producers and argued that such a phenotype could be caused by increased transcriptional load

and expanded ER associated with secretory protein production. By simulating the theoretical metabolite demand using a network model, they linked the metabolic changes in protein producing cells to increased C1-unit, nucleotide and lipid metabolism, which led to specific adaptations in the amino acid metabolism and increased secretion of glycine and glutamate. The authors concluded that C1 and lipid metabolism seem important targets for improvement of protein production in mammalian cells.

The glycoprotein hormone erythropoietin (EPO) is a commonly used model protein in development of CHO-based bioprocesses (Yoon et al. 2005; Surabattula et al. 2011; Choi et al. 2007; Sung et al. 2004) and metabolic engineering of CHO cells for improved protein production (Kim et al. 2011; Kim et al. 2004). The typical EPO expression levels from clones with no gene amplification are reported in the range of 1-10 pg/cell/day (Zhou et al. 2010; Kim & Lee 2009; Yoon et al. 2003), which is substantially lower than e.g. antibody production processes.

The aim of the current study was to discover bottlenecks in EPO production in CHO cells and characterize the burden of heterologous protein production under growth dependent and independent conditions. For this, a panel of stably EPO expressing CHO-K1 clones spanning a 25-fold productivity range was established and characterized in batch and chemostat cultures. We employed a multi-omic physiological characterization including NMR-based metabolic footprinting (exo-metabolome) of sugars, organic acids and amino acids, LC-MS based metabolite fingerprinting (endo-metabolome) of glycolytic intermediates, NAD(P)H/NAD(P)$^+$ and ANP's. Quantitative PCR (qPCR), quantitative reverse transcription PCR (qRT-PCR), western blots (WB) and Affymetrix CHO microarrays were used to assess EPO gene copy numbers, EPO gene expression, intracellular protein

levels and genome-wide gene expression analysis of differentially expressed genes functionally related to secretory protein processing, respectively.

Materials and methods

Cell lines and media

The EPO-expressing cell lines were developed from the ATCC (Manassas, Virginia) CHO-K1 line cat no. CCL-61. Prior to cell line development the parental cell line was adapted for suspension and serum-free culture in a complex animal-component free Novo Nordisk proprietary medium supplemented with 4 mM L-glutamine (Thermo Scientific, Waltham, MA). During development of EPO-expressing cell lines the media were supplemented with 2.5 mL anti clumping agent (Gibco) per 1 L medium and Penicillin-Streptomycin mix (Gibco) in concentrations of 100 unit/mL of penicillin and 100 µg/mL streptomycin. 600 µg/mL Geneticin / G418 (Gibco) was applied as selection pressure one day after transfection and throughout the cell line generation process.

Primers

Primers for Uracil-specific excision reagent (USER) cloning procedure (Table I) were designed according to the USER cloning design scheme in (Lund et al., 2014) and purchased from Integrated DNA Technologies (Leuven, Belgium). Primers for specific amplification of target sequences in *hEPO*, β-actin (*Actb)* and glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) (Table I) were designed using the online quantitative PCR primer design tool from Roche, which is based on the Primer3 software (Untergasser et al. 2012) and gene sequences were retrieved from www.chogenome.org (Hammond et al. 2012).

Chemicals for analysis of intercellular metabolites

Isotope-labeled standards were purchased from Silantes Gmbh (München, Germany). All other standards of metabolites were obtained from Sigma-Aldrich (St. Louis, MO), except for acetyl coenzyme A that was produced by Santa Cruz Biotechnology (Dallas, TX). High purity solvents and reagents were used in order to reduce the background noise from impurities as much as possible. The solvents acetonitrile and methanol used for extraction were HPLC grade from Sigma-Aldrich while the methanol used for chromatography was LC-MS grade from Fluka. All water was milli-Q purified. The ion pair reagent tributylamine (TBA) (HPCL grade) was from Sigma-Aldrich, while the acetic acid (LC-MS grade) was from Fluka.

Plasmid construction

A vector plasmid pEPO-NEOR was assembled using the uracil-specific excision reagent (USER) based FAST-mediated vector assembly procedure as previously described (Lund et al. 2014). Neomycin resistance was included in the construct as selection marker. The human erythropoietin (EPO) gene (Powell & Berkner 1986) in the plasmid construct was codon-optimized for CHO and synthesized from Geneart (Regensburg, Germany). The mammalian expression vector pU0002 (Hansen et al. 2011) harboring an *E. coli* origin of replication element and an ampicillin resistance gene was used as plasmid backbone. The EPO gene was under control of the human cytomegalovirus (CMV) promoter and flanked by the bovine growth hormone polyadenylation signal (BGHpA), while the NEO^R gene was regulated by the simian vacuolating virus 40 (SV40) promoter and polyadenylation signal (SV40pA). USER elements harboring promoter regions, polyadenylation signals, the NEO^R gene, and the protein backbone were produced exploiting PCR primers and protocols from (Lund et al. 2014). Analogously, a USER element with EPO was prepared using the uracil-containing primers found in table I. The NEO^R gene was assembled with its promoter and

polyadenylation signal in one USER cloning event exploiting the USER enzyme mix (New England Biolabs, Ipswich, MA) and the competent *E. coli* DH5α strain (Invitrogen, Carlsbad, CA) as described in details in (Lund et al. 2014). Subsequently, the formed selection marker element was amplified by PCR and used in a second USER cloning procedure for generation of the vector plasmid pEPO-NEOR. Plasmid sequence was verified by sequencing (Star SEQ, Mainz, Germany).

Generation of EPO-expressing cell lines

Transfection of the parental CHO-K1 cell line with the plasmid vector pEPO-NEO^R was performed by electroporation in a BioRad GenePulser Xcell set to deliver a single pulse of 900 µF at 300 V and infinity resistance in a 4 mm cuvette. As positive control a subset of cells were transfected with a mammalian expression vector with the gene for enhanced green fluorescent protein (eGFP) and neomycin resistance. The control transfection was used to estimate transfection efficiency, follow cell death, clone expansion, and transgene expression. Prior to each transfection 40 µg of plasmid DNA was added directly to the cuvette containing 10^7 cells in growth medium. Twenty-four hours after transfection G418 selection pressure was added and the transfected cells were split into two. Single clones were isolated from one half of the transfected cells in a limiting dilution experiment with twenty 96-well plates containing either 500 or 1000 transfected cells/well. During two weeks of cultivation one 96-well plate was exposed to microscope inspection daily to observe initial cell death and stable clones expanding. From the untouched 96-well plates circular monoclonal cultures were screened for EPO production using a dot blot procedure followed by WB and enzyme-linked immunosorbent assay (ELISA) (see below) and expanded further.

The second half of the transfected cells were maintained as a polyclonal shake flask culture for three weeks. For the first two weeks the culture volume was gradually decreased

in each passage to maintain a viable cell density of $0.3 \cdot 10^6$ cells/ml. Single clones were isolated from the polyclonal culture by limiting dilution into 384-well plates and robot-assisted single clone selection in a Cello system (TAP Biosystems, Royston, UK). The cells were cultivated and photos were taken for 13 days with medium change every 6 days. Single clone cultures were screening for EPO production and scaled up to 30 ml shake flask cultures.

Screening cell lines for EPO production

Isolated monoclonal cell lines were screened for EPO production using WB and selected clonal cultures were up-scaled and evaluated further using the Quantikine IVD ELISA kit (R&D systems, Minneapolis, MN) following the manufacturer's protocol.

The Invitrogen NuPAGE system was used for WB. Samples of culture supernatant were drawn and centrifuged at 15000 x g for 1 min and treated following the NuPage guidelines for preparation of reduced samples and peptide N-glycosidase treated samples using PNGases F (New England Biolabs). Samples were run at 12% NuPAGE Novex bis-tris mini gels with MOPS running buffer in an Xcell SureLock mini cell at 200 V (constant) for 45 min with MagicMark™ XP Western protein standard (Invitrogen) and Full-range rainbow molecular weight marker (GE Healthcare). Gel separated proteins were transferred by an Invitrogen iBlot device to a nitrocellulose membrane with 0.45 µm pore size (Invitrogen). 2.0 % TBS-T, was used as blocking buffer and for washing steps 0.5 % TBS-T was employed. The membrane was incubated with 1 µg/ml polyclonal rabbit anti-EPO antibody (AbCam, Cambridge, United Kingdom) in 10 ml 0.5 % TBS-T at room temperature with gentle shaking at 45 rpm for 45 min. Following three washing steps with 0.5 % TBS-T the membrane was incubated with 0.2 µg/ml IRDye 680 goat anti-rabbit (Li-Cor Biosciences) fluorescent labeled secondary antibody in 0.5 % TBS-T for 45 min with shaking at 45 rpm.

The membrane was analyzed in a Li-Cor Odyssey infrared imaging system. Supernatant from eGFP clones served as negative control.

Cell culture

Cell culture was performed in vented Erlenmeyer shake flasks (Corning, NY) in a shaking incubator operated at 36.5°C, 5 % CO_2 and 140 rpm. Cells were cultured in repeated batch cultivation during the development of EPO-expressing cell lines. The cells were passaged twice a week and the viable cell density was adjusted to 0.3 x 10^6 cells/mL.

Pre-cultures were initiated from frozen cells and cultivated as above, but without selection pressure. The pre-cultures were passaged every other day to ensure growth at maximum specific growth rate.

Bioreactor cultivation and analysis

Parental and recombinant CHO-K1 cells were cultivated in 1.5 L bioreactors (Eppendorf DASGIP multi-fermentor system, Jülich, Germany) with a working volume of 1 L. Temperature was maintained at 36.5°C with an agitation rate of 200 rpm using two three-way segmented impellers. Dissolved oxygen was maintained at 50 % of air saturation using air, O_2 and CO_2 operated at a constant flow rate of 0.6 L/h. Culture pH was maintained at 7.15 with a deadband of 0.25 using intermittent CO_2 addition to the gas mix and 2M sodium carbonate. Culture pH and pO_2 was measured on-line and calibrated to an offline reference RAPIDpoint 500 blood gas analyzer (Siemens Healthcare Diagnostics, Erlangen, Germany) subsequent to inoculation. Cell number, viability, cell size and aggregation was measured using a CedeX HiRes (Roche, Basel, Switzerland), extracellular concentrations of glucose, lactate, glutamine, glutamate and ammonium was measured with a Bioprofile 100^{PLUS} (Nova

Medical, Waltham, MA). Supernatant samples for extracellular EPO quantitation were stored at -80°C until HPLC analysis.

Batch cultures were seeded with 0.3 x 10^6 cells/mL and samples were drawn on a daily basis and analyzed for cell density, viability, cell size and aggregation rate. Extracted culture supernatants were analyzed for glucose, lactate, glutamine, glutamate, ammonium, EPO, pH, pO_2 and pCO_2. Genomic DNA was extracted after 48 hours and analyzed for EPO gene copy numbers by quantitative PCR. The culture was terminated after 160 hours.

Chemostat cultures were seeded with 0.3 x 10^6 cells/mL and chemostat cultivation mode was initiated 72 hours subsequent to inoculation with a constant dilution rate of 0.3 volumes per day. The cultures were sampled daily and analyzed for cell density, viability, cell size and aggregation rate. The supernatant was analyzed for glucose, lactate, glutamine, glutamate, ammonium, EPO, pH, pO_2 and pCO_2. Samples for metabolic footprinting were analyzed for amino acids, sugars and organic acids by quantitative NMR analysis (Spinnovation Biologics, Nijmegen, Netherlands). Genomic DNA and RNA was extracted and analyzed for EPO gene copy number and EPO gene expression level by qPCR and qRT-PCR, respectively. Selected cultures were subjected to microarray based gene expression analysis.

HPLC quantitation of erythropoietin

EPO from thawed supernatant samples was quantified by RP-HPLC on an Agilent 1200 using an XBridge C8 4.6 x 150 mm (3.5 µm) column (Waters), operated at 42°C and a flow rate of 1 mL/min. Buffer A was composed of 0.1 % TFA in milliQ water and buffer B was composed of 0.07 % TFA in acetonitrile. The elution gradient consisted of 30-70 % buffer B over 16 min. Protein detection was performed by UV light absorption at 214 nm and EPO

concentration was determined using human erythropoietin (Cell Signaling, Danvers, MA) as standard.

Preparation of DNA, RNA and cDNA

Genomic DNA (gDNA) was isolated from pellets of 3 x 10^6 CHO cells using a DNAeasy blood and tissue genomic DNA purification kit (Qiagen, Hilden, Germany) following the manufacturers instructions. DNA concentration and purity was determined using a Nanodrop 8000 (Thermo Scientific, Wilmington, DE, USA). Samples with $A_{260/280}$ ratios ≥ 2 were considered to be of sufficient purity.

For total RNA isolation, 3 mL culture sample was extracted and centrifuged at 900 x g for 5 min. The supernatant was discarded and the cell pellet was homogenized in 2 mL Trizol reagent (Invitrogen) and stored at -80°C. Total RNA was extracted using an RNA plus mini kit (Qiagen) according to the manufacturer's instructions including column-based digestion of DNA. Total RNA quantity was determined spectrophotometrically using a Nanodrop 8000 (Thermo Scientific, Wilmington, DE) and RNA sample integrity was determined using an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA) ensuring RIN values above 9.0.

cDNA was generated from total RNA using a High Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster city, CA) according to the manufacturers instructions.

Determination of relative *hEPO* gene copy numbers and mRNA levels

Relative *EPO* transgene copy numbers and mRNA levels were determined using real-time quantitative PCR on gDNA and mRNA, respectively. Primer pairs were tested for specificity and amplification efficiency. Primer dimerization and specificity was investigated using

melting curve analysis, which revealed a single thermal transition confirming that the primers were specific for the target genes and indicating absence of primer dimerization. Standard curves were generated from serial dilutions of pooled gDNA in triplicates and amplification efficiencies close to 100 % were achieved for all primer pairs. Primers targeting the commonly used reference genes *Gapdh* and *Actb* were screened for amplification efficiency and *Gapdh* was selected as reference gene as the primers produced amplification efficiencies closer to 100 %. Quantitative PCR was performed using a QuantiFast SYBR Green PCR Kit (Qiagen) containing the fluorescent dye SYBR green I and ROX as fluorescent reporter. Quantification of relative *EPO* gene dosage and expression level was carried out in 384 well plates in a 7900HT FAST Real-Time PCR System (Applied Biosystems) with a reaction volume of 10 µL. All PCR reactions were run in triplicates. The assay was executed with the following thermal profile: 10 min heat activation of the polymerase at 95°C followed by 40 amplification cycles consisting of DNA dissociation at 95°C for 5 s and primer annealing at 60°C for 20 s. The dissociation stage consisted of a linear temperature ramp from 60°C to 95°C over the course of 10 min. The C_T values were computed using the AutoC_T algorithm found in the software package SDS 2.4 (Applied Biosystems). For calculation of gene copy numbers cells were assumed to be diploid.

Transcriptomics sample preparation and data analysis

Chemostat cultivations of three clones (clone 1, clone 4 and clone 7) were carried out in two parallel cultures (biological replicates) and samples for RNA isolation were taken during the steady state phase of each culture, as determined by constant concentrations of medium components (amino acids and sugars). RNA samples were isolated from the culture as described above. RNA sample integrity was determined using Agilent 2100 Bioanalyzer and RNA 6000 Nano LabChip kit (Agilent, Santa Clara, CA), ensuring RIN values above 9.0, and

total RNA quantity was determined with a NanoDrop 3300 UV–Vis spectrophotometer (Thermo Scientific, Rockford, IL). Using the GeneChip Hybridization, Wash and Stain Kit, the probe preparation and hybridization to Affymetrix CHO Gene 2.0 ST Arrays were performed according to manufacturer's instructions (Affymetrix GeneChip Expression Analysis). Washing and Staining of arrays were performed using the GeneChip Fluidics Station 450 and scanning with the Affymetrix GeneArray 3000 7G Scanner (Affymetrix, Santa Clara, CA). The Affymetrix GeneChip Command Console Software (AGCC) was used to generate CEL files of the scanned arrays.

Differential gene expression analysis was performed using the Transcriptome Analysis Console (TAC) 2.0 (Affymetrix) software package using One-Way ANOVA, p-values were corrected for multiple comparisons by Benjamini & Hochberg False Discovery Rate (FDR). Transcripts with a FDR p-value <0.05 were considered statistically significant.

Western blot analysis of intracellular EPO retention

Intracellular EPO retention was examined using SDS-PAGE in conjunction with WB analysis. Intracellular proteins were extracted from pellets of 5 x 10^6 cells in mid-exponential phase using 1 mL Mammalian Protein Extraction Reagent with complete™ protease inhibitor cocktail added (Thermo Scientific). The mixture was left to react for 10 minutes with gentle shaking and cell debris were removed by centrifugation at 14000 x g for 15 minutes. For electrophoresis, 28 µL total protein sample was denatured with 4µL NuPage Sample Reducing Agent (Invitrogen) and 8 µL NuPage LDS Sample Buffer (Invitrogen) at 80°C for 5 minutes and size fractionated on a 12 % NuPAGE Novex Bis-Tris mini gel with MOPS running buffer. Gel separated proteins were transferred to a 0.45 µm pore size nitrocellulose membrane using an iBlot (Invitrogen), mouse anti-EPO (RnD Systems, Minneapolis, MN) was used as primary antibody and a fluorescent labeled donkey anti-mouse antibody (Licor)

was used as secondary antibody. The fluorescence was quantified using an Odyssey CLx (Licor) with human erythropoietin (Cell Signaling) as positive control.

Quenching and extraction of intracellular metabolites

For analysis of intracellular metabolite pools, 10^7 cells were extracted from mid-exponential batch cultures and immediately quenched with four sample volumes 0°C 0.9 % w/v NaCl on ice (inspired by Dietmair et al. 2010). The cooled cell suspension was immediately spun down at 1000 x g for 1 min at 0°C and the supernatant discarded. 1mL of -79°C methanol was added to the cell pellet followed by addition of an internal standard mixture containing 10 µg/mL of [U-^{13}C] ATP and [U-^{13}N] AMP and flash freezing in liquid nitrogen (inspired by Sellick et al. 2010). Samples were stored at -80°C before thawing on ice and two successive extractions were performed with 1 mL 50 % v/v acetonitrile in water (inspired by Dietmair et al. 2010). The extraction procedure included addition of solvent solution, resuspension of cell pellet by vortexing, incubation on ice for 10 min and separation of cell debris and liquid phase by centrifugation at 4200 x g for 5 min. The pooled extraction supernatants were filtered through a 0.45 µm teflon syringe filter Ø17 mm (National Scientific, Rockwood, TN). 8 mL acetonitrile was added to the filtrate to facilitate water evaporation before drying under nitrogen atmosphere at room temperature. The extracted metabolites were dissolved in 150 µL milliQ water containing 10 mM tributylamine and 10 mM acetic acid resulting in a final concentration of 1 µg/mL of each of the internal isotope-labeled standards. Prior to the analysis the samples were filtrated using a 0.45 µm teflon syringe filter Ø17 mm (National Scientific).

Ion-pair liquid chromatography tandem mass spectrometry

All LC-MS/MS experiments were performed on an Agilent 1290 Infinity LC coupled with an

Agilent 6460 triple quadrupole MS analyser equipped with electrospray ionization source. The MS was operated in negative multiple reaction monitoring (MRM) mode.

10 µg/mL single standard solutions in 10 mM TBA and 10 mM acetic acid were used to optimize the compound specific MS and ion source parameters. The two most intense MRM transitions for each metabolite were determined in a direct infusion experiment using a KDS100 infusion pump with a flow rate of 9,8 µL/min. Then, for each chosen MRM transition the collision energy (CE), fragmentor and cell accelerator voltages (CAV) were optimized by injecting 1 µl of 10 µg/mL single standard solutions. When investigating the optimal compound specific parameters for the MS/MS analysis, the following range of voltages were tested: fragmentor voltage 90-130 V in steps of 10 V, CE 5-35 V in steps of 5 V and CAV 3 and 4 V. The MRM's used for the analysis are given in supplementary materials. The best compound specific parameters were those giving the most intense LC-MS peak. The ion source dependent parameters were as follows: gas temperature 300°C; sheath gas temperature 400°C; nebulizer gas flow rate 8 L/min; nebulizer pressure 50 psi; and capillary voltage 4500 V. Nitrogen was used as collision gas. The entrance potential (ΔEMV) and dwell time were kept at 500 and 30 ms respectively for all transitions.

The chromatographic separation was obtained on a Luna 2.5µ C18(2)-HST (100 x 2.0 mm) HPLC column (Phenomenex, Aschaffenburg, Germany) operated at 40°C. Eluent A was water containing 10 mM tributylamine and 10mM acetic acid and eluent B was 90% (v/v) methanol containing 10 mM tributylamine and 10mM acetic acid. The gradient was stepwise 0-5 min, 0% B; 5-10 min, 0-2 % B; 10-11 min, 2-9 % B; 11-16 min, 9% B; 16-24 min, 9-50% B; 24-28 min, 50% B; 28-28.5 min, 50-100% B; 28.5-30 min, 100% B; 30-30.5, 100-0% B; 30.5-36 min, 0%. The final 5.5 min were used for equilibration of the column prior to the next run. The flow rate was 0.3 mL/min and the sample injection volume was set to 5 µL.

1 mg/mL single standard stock solutions in water were used to prepare 10 µg/mL mixture of the compounds of interest in eluent A. The latter mixture was used to prepare the calibration solutions with concentrations ranging from 0.05 to 10 µg/mL. Standard curves used for the quantification were constructed by plotting the peak area of the compounds against the concentration. For the compounds for which internal standards were available the calibration curves were constructed by plotting the ratio of the peak area of labeled and unlabeled compounds against their concentrations. A chromatogram of detected compounds in mammalian cell extracts is supplied in supplementary materials.

Metabolic network reconstruction

A draft network reconstruction of the glycolytic and amino acid catabolic pathways in CHO cells was generated using the mouse metabolic pathways as template. Biochemical pathway data from mouse metabolism was retrieved from the Kyoto Encyclopedia of Genes and Genomes database (Kanehisa & Goto 2000; Kanehisa et al. 2014) and homologous gene sequences in the CHO genome were identified using the Chinese hamster genome database www.CHOgenome.org (Hammond et al. 2012). The draft network reconstruction was further refined by careful curation of gene-protein-reaction relationships using manual genome annotation and literature evidence. The finalized reconstruction featured 319 proteins catalyzing 183 reactions with 188 metabolites (metabolic map is supplied in supplementary materials).

Statistical analysis

The statistical test for determination of physiological differences between clone populations was performed using Student's t-test with a significance level of $\alpha = 0.05$.

Results

Cell line generation and clone selection

Seven single cell clones were selected based on proliferation rate and EPO expression to establish a panel of stable clones with specific EPO productivities (q_{EPO}) ranging from less than 0.2 to 5 pg/cell/day (determined in exponential growth phase), thus covering a 25-fold range of productivity (Figure 1). All EPO producing clones and a non-transfected parental clone were adapted to the growth medium (Q-CM105) to exclude the influence of ongoing medium adaptation on physiological characterization. During the adaptation phase, clones were monitored for specific growth rate, specific glucose and glutamine uptake rates, specific lactate and ammonium secretion rates and specific EPO production rate. After 20 days in 100 mL repeated batch culture, all measured parameters had stabilized (i.e. remained within 7 %). Thus, the clones were considered fully adapted to the growth medium and a master cell bank was established.

EPO production has no effect growth, nutrient uptake or by-product secretion

Clones C1-7 and the parental clone were physiologically characterized in duplicate batch cultivations in bioreactors under nutrient excess conditions, to ensure maximum specific growth rate (supplementary materials). To assess the physiological impact of EPO production, the control was compared to the EPO producing clones (data displayed in table II). No significant difference was found in growth characteristics (i.e. specific growth rate and biomass yield), excluding major physiological stress from EPO production. An analysis

of correlation (supplementary materials) between the cell phenotypic variation displayed in Table II and q_{EPO} was performed to identify patterns in clone physiology that might explain the difference in q_{EPO}. The analysis identified no correlations in the dataset (Table II), suggesting that an in-depth physiological analysis was required to discover phenotypic markers for high q_{EPO}.

Comparison across EPO producing clones reveals no detrimental effect on glucose metabolism

In order to assess possible metabolic impact of differential EPO expression on energy metabolism, we performed a quantitative characterization of intracellular metabolites related to glucose energy and redox metabolism (i.e. specific glycolytic intermediates, NAD(P)H-/NAD(P)$^+$ and ANP's). For this, triplicate batch cultures of all EPO producing clones were sampled in parallel, during mid-exponential growth phase and metabolite profiles were generated using an extraction technique that does not differentiate between cellular compartments, thus picturing the average concentration of intracellular metabolites. The differences in concentrations of adenosine phosphates and nicotinamide adenine dinucleotides did not exhibit a marked correlation to q_{EPO} (linear regression analysis, $R^2 < 0.36$) (Figure 2 A+C). The adenylate energy charge (AEC) ratio represents the amount of metabolically available energy stored in the adenine nucleotide pool (Atkinson 1968). The catabolic- and anabolic reduction charges represent the redox state of the cell (Andersen & von Meyenburg 1977). The observed distributions (Figure 2 B, D, E) indicate that EPO production is not limited by insufficient energy availability from adenosine phosphates or nicotinamide adenine nucleotides (linear regression analysis, $R^2 < 0.36$).

Furthermore, intracellular concentrations of several carbon metabolites from glycolysis and acetyl coenzyme A were determined and compared between clones (Figure 3).

As observed for adenosine phosphates and nicotinamide adenine nucleotides, the differential q_{EPO} was not reflected in metabolite concentrations indicating that EPO production is not limited by glucose metabolism (linear regression analysis, $R^2 < 0.33$).

Chemostat cultivation of three EPO producing clones show temporal correlations in gene expression and EPO titer

To identify the bottleneck in the protein production pathway, three clones (C1, C4 and C7) were selected for an in-depth physiological characterization under growth-limited conditions in duplicate chemostat cultivations. The cultures were continued for 31 days with a fixed specific growth rate of 0.3 day^{-1} corresponding to 15 generations at 30 % of maximum growth rate. The chemostat cultivation mode was selected to normalize for growth-related effects on protein productivity across the three clones, thus unveiling physiological variation in protein production efficiency regardless of maximum growth capacity. The assumption here was that normalization of the specific growth rate lead to normalization of metabolic fluxes and therefore picture the intrinsic metabolic efficiency of protein production between the clones. Samples were taken daily from each chemostat culture and the secreted protein levels, EPO gene copy numbers, mRNA levels (Figure 4) and amount of intracellular accumulated EPO were determined (supplementary materials). The viable cell density (Figure 4 A), stabilized at approximately 5 million cells per mL after 10 days. Analysis of spent growth medium suggested that the cultures reached steady-state at day 12, as concentrations of medium components (amino acids and sugars) and metabolic by-products (lactate and ammonium) were constant in all cultures from this time-point (supplementary materials). The dynamics of EPO titers pictures three distinct phases (Figure 4 B). In phase I (day 1-10) the EPO titers decrease as the cells adjust to the imposed growth limitation and the steady state. During phase II (day 10-20) the cells reach steady-state and protein titers are

relatively stable in all cultures. In phase III (day 20-31) the EPO titers increase corresponding to an increase of q_{EPO} by 56 %, 74 % and 83 % for clone 1, clone 4 and clone 7, respectively in phase III relative to phase II (Figure 4, E+F bars). The EPO gene copy numbers were determined by qPCR using relative quantitation with *Gapdh* as reference gene. The dynamics of EPO gene copy numbers feature a steady increase over the course of the cultivation (Figure 4 C). Starting with 1.5 relative EPO gene copies, the determined gene copy numbers slowly increase towards 2 EPO gene copies, suggesting a culture-average absolute EPO gene copy number of 3 at the beginning of the cultivation and 4 in the end. The dynamics of EPO gene expression pictured a decrease around day 12 consistent in all cultures (Figure 4 D). The basis of the sudden decrease is unknown, but the timing correlates with depletion of lactate in the growth medium. From day 20, the EPO gene expression increased in all clones throughout the cultivation, correlating well with the increased EPO titers in phase III.

Post-transcriptional protein processing efficiency of EPO in the protein secretory pathway correlates with specific EPO productivity across clones

For determination of differences in EPO transcription efficiency across clones, we compared the ratios of culture-average EPO gene expression per EPO gene, i.e. the ratio of EPO mRNA to EPO gene (Figure 4 E). It was noticed that the transcriptional efficiency of clone 4 and clone 7 was identical throughout the experiment and that the transcriptional efficiency of clone 1 was consistently 20 % lower than the other clones. To determine differences in post-transcriptional processing of EPO across the clones, we compared the culture-average ratios of EPO titer and EPO gene expression, i.e. EPO titer per EPO mRNA, thus reflecting the efficiency of protein translation and secretory protein processing (protein folding, -maturation and -secretion) across clones (Figure 4 F). It was noticed that the post-transcriptional

efficiency was significantly higher in clone 7 relative to clone 4 and clone 1 and corresponded well to the observed difference in q_{EPO}. Therefore, we investigated whether different amounts of EPO were retained intracellular in the clones. For this, total cellular protein extracts were separated using SDS-PAGE and analyzed for EPO contents using WB (supplementary materials). The differences in intracellular EPO levels corresponded to the observed extracellular EPO titers (Figure 4 B) indicating that EPO is not retained intracellular in any clones.

Global gene expression analysis indicate adaptation of gene expression levels of amino acid catabolic genes to preserve most abundant amino acids in EPO

To identify differentially expressed genes functionally related to secretory protein processing across the EPO producers we performed a global gene expression analysis comparing the highest and lowest EPO producers (clone 7 and clone 1, respectively) during the steady-state phase of chemostat culture in phase II (triplicate samples were generated from day 12, 15 and 18). The differential gene expression analysis identified no enrichment in the gene expression landscape of genes related to protein translocation, protein folding, protein glycosylation or vesicular transport (supplementary materials), indicating that neither of these processes was limiting the protein productivity. Next, we investigated whether the protein production bottleneck was reflected in differential expression of metabolic genes. For this analysis, we generated a network reconstruction of the glycolytic pathway and the amino acid catabolic pathways, as these are the most active catabolic pathways and thus most likely to limit energy metabolism (the reconstructed metabolic network is displayed in supplementary materials). The network reconstruction served as a framework for meaningful interpretation of the differential gene expression data on a pathway level. The results indicated a general up-

regulation of glycolytic genes in clone 7, suggesting a possible increased energy demand in this clone. Interestingly, when inspecting the differential gene expression levels of amino acid catabolic genes, we discovered a tendency towards preservation of the most abundant amino acids in EPO in the high producer relative to the low producer (i.e. decreased transcription level of genes responsible for degradation of the amino acids most frequently found in EPO) (Figure 5). Specifically, 12 of the 13 most abundant and non-secreted amino acids in EPO had reduced expression of catabolic reactions in the high producer relative to the low producer (Figure 5B). Thus, the result indicated possible regulatory adaptation of gene expression towards decreased amino acid catabolism specific for the most abundant amino acids in EPO, in the high producer relative to the low producer. It was noticed that the observation was followed by an increase of q_{EPO} by 56 % and 83 % in the clone 1 and clone 7, respectively (phase III, Figure 4 B).

Discussion

Comparison across EPO producing clones revealed no apparent bottlenecks in the protein expression and secretory pathway or energy metabolism

The secretory production of proteins in CHO cells can be characterized as a cascade of protein modification and quality control steps catalyzing the post-translational processing of a nascent polypeptide into a functionally mature protein (Hussain et al. 2014). The overproduction of a heterologous protein increases the trafficking through the secretory pathway to the limit of the protein processing capacity leading to productivity bottlenecks. To increase the knowledge of the bottleneck associated with secretory production of EPO in CHO cells, we established a panel of CHO-K1 clones spanning a 25-fold range of specific EPO productivity and assessed the phenotypical differences at multiple stages within the protein expression and secretion pathway.

The comparison of transcriptional efficiency (Figure 4, E) showed a lower transcription rate per EPO gene in clone 1 compared to clone 4 and clone 7 throughout the experiment, indicating that the EPO gene was inserted in a locus with less transcriptional activity in clone 1. While clone 4 and clone 7 showed identical transcriptional efficiencies, the comparison of post-transcriptional efficiency (Figure 4, F) revealed that clone 1 and clone 4 were severely limited in EPO secretion per EPO transcript compared to clone 7 (23 % and 50 % of C7 at day 15, respectively). It was observed that the difference in post-transcriptional efficiency corresponded to the difference in q_{EPO} indicating that the expression bottleneck was enrooted downstream of transcription (i.e. translation, translocation, protein folding, -glycosylation and -transport). The differential EPO expression was not reflected in intracellular protein concentration as determined by Western blot, as this correlated well with the difference in extracellular protein concentration, indicating that post-translational processing of EPO in the secretory pathway is not a bottleneck. This indication was underlined by the fact that the global gene expressing analysis of clone 1 and clone 7 found no significant (p-values > 0.05) difference in expression level of single genes or expression enrichment within a group of genes functionally related to secretory protein production (i.e. genes involved in translocation, protein folding, -glycosylation and -transport).

The determination of gene- and transcript levels during prolonged chemostat cultivation led to some noteworthy observations. The slightly increasing trend of EPO gene copy numbers was surprising. However, the effect may be explained by presence of a sub-population of cells with different copy numbers of *hEPO* or *Gapdh*, as previously demonstrated by Beckmann et al. 2012. Similarly, the sudden decrease of EPO transcripts around day 12 (Figure 4, D) was surprising. The basis of the decrease was unknown, but the timing in all 5 cultures correlated well with the depletion of lactate in the growth medium and may be associated with a metabolic shift.

It was investigated whether the differential EPO expression across the clones was caused by a bottleneck in carbon and/or energy metabolism. For this, intracellular metabolites were sampled in mid-exponential growth phase as this was assumed to picture the maximum metabolic capability of each clone. Comparison of intracellular concentrations of adenosine phosphates and nicotinamide adenine dinucleotides across clones showed no correlation to q_{EPO} (Figure 2). This observation indicated that the energy metabolism was keeping up with the increased energy requirement in the EPO producing clones, which is in agreement with similar studies of other mammalian cell types (Khoo et al. 2007; Niklas et al. 2013). Furthermore, the lack of correlation between concentrations of glycolytic intermediates and q_{EPO} (Figure 3) indicated that glucose metabolism was not limiting for EPO productivity in batch culture. However, in the chemostat culture, we observed a change in the expression landscape of metabolic genes between the two EPO producing clones. The genes in the glycolytic pathway were generally up-regulated in the high producing clone, possibly reflecting an increased energy demand corresponding to the increased EPO productivity of this clone. That is, the normalization of growth rates in chemostat culture normalized the metabolic energy consumption from growth, thus allowing the quantification of energy requirement from heterologous protein production. Increased glycolytic flux in response to protein production during glucose-limited growth-restricted culture has been demonstrated in the eukaryotic production host *P. pastoris* (Heyland et al. 2010).

Heterologous protein production causes metabolic changes in favor of the produced protein

Heterologous protein production imposes a metabolic burden on the host cell metabolism, which causes redistribution of metabolic precursor fluxes to meet the increased anabolic demand for e.g. nucleotides for synthesis of RNA and activated sugar precursors associated

with secretory protein production (Niklas et al. 2013). The same study demonstrated that anabolic demand for nucleotide biosynthesis results in extracellular secretion of glycine and glutamate. Interestingly, we found that during steady state in chemostat culture, extracellular concentrations of glycine and glutamate were 1.8-fold and 2-fold higher in C7 relative to C1, respectively. This indicated that the secretion rates of glycine and glutamate increased with q_{EPO}, suggesting that the findings of Niklas et al. (2013) in human cells expressing α_1-antitrypsin are also valid for CHO cells expressing EPO.

The use of a nutrient-limited cultivation format restricts the possibility to increase nutrient uptake and inflict regulatory changes on cell metabolism, which may lead to flux-redistribution in favor of the heterologous protein. To further increase the knowledge on the adaptability of CHO cell metabolism, we performed a comparative transcriptome analysis of two clones with 25-fold differential EPO productivity in glucose-limited chemostat cultivations at D = 0.3 day^{-1}. Interestingly, we observed a change in the gene expression landscape of catabolic genes between the clones. The genes in the glycolytic pathway generally showed higher expression levels in the high producing clone, possibly reflecting an increased energy demand corresponding to the increased EPO productivity of this clone. Furthermore, the comparison of gene expression levels in the amino acid catabolism revealed a regulatory change around the amino acids, which are most abundant in EPO and not secreted from the cell. That is, the gene expression levels of enzymes producing these amino acids were generally up-regulated and expression levels of enzymes consuming the same amino acids were generally down-regulated in the high producer relative to the low producer (Figure 5). This observation indicated a comparatively larger degree of metabolic adaptation to EPO production in the high producer, which may explain the larger increase of q_{EPO} in the high producer in phase III of chemostat culture (83 % vs. 56 % in high- and low producers, respectively). Based on these data, we speculate that the amino acid metabolism in CHO cells

may undergo adaptation in favor of the produced heterologous protein during long-term cultivation. The adaptation of gene expression levels in amino acid metabolism in favor of heterologous protein production during prolonged chemostat cultivation has been reported before in the eukaryotic protein production host *S. cerevisiae* (Kazemi Seresht et al. 2013).

In conclusion, we provide evidence that EPO production up to 5 pg/cell/day is not limited by metabolism (i.e. glycolysis and associated energy metabolites) or bottlenecks in gene dosage, transcription and post-translational processing of EPO. Furthermore, we showed that glutamate and glycine secretion is increased in the high producing EPO clone, relative to the low producing clone, echoing the findings of Niklas et al. (2013) thus indicating possible anabolic demand for nucleotides and lipids, which could be candidate targets for medium supplementation to improve protein productivity.

Finally, we demonstrate that heterologous protein production can inflict metabolic changes in favor of the produced protein during prolonged chemostat cultivation. The observed adaptations of glycolysis and amino acid metabolism were followed by increased protein productivity in phase III (83 % vs. 56 % in high- and low producers, respectively), suggesting that metabolic engineering of amino acid metabolism to reduce catabolism of amino acids present in the target protein could improve specific protein productivity in continuous culture. It was not possible to verify the reduced amino acid catabolism at the metabolite level using metabolic foot printing, thus future work should include quantification of intracellular levels of amino acid catabolic proteins or metabolic flux analysis to verify the suggested link between amino acid catabolism and heterologous protein production in chemostat culture.

Acknowledgements

We would like to thank Carsten Leisted, Jens Jacob Hansen and Anja Kallesøe Pedersen for support with bioreactor cell culture experiments, cell line development and development of the HPLC-based EPO quantitation assay, respectively.

H.F.K thanks the Novo Nordisk Foundation and the Lundbeck Foundation for financial support.

The authors declare no conflict on interest.

References

Andersen, klaus B. & von Meyenburg, K., 1977. Charges of nicotinamide adenine nucleotides and adenylate energy charge as regulatory parameters of the metabolism in Escerichia coli. *Journal of Biological Chemistry*, 252(12), pp.4151–4156.

Atkinson, D.E., 1968. Energy charge of the adenylate pool as a regulatory parameter. Interaction with feedback modifiers. *Biochemistry*, 7(11), pp.4030–4034.

Beckmann, T.F. et al., 2012. Effects of high passage cultivation on CHO cells: a global analysis. *Applied microbiology and biotechnology*, 94(3), pp.659–71.

Bolt, G., Kristensen, C. & Steenstrup, T.D., 2008. More than one intracellular processing bottleneck delay the secretion of coagulation factor VII. *Thrombosis and Haemostasis*, pp.204–210.

Borth, N. et al., 2005. Effect of increased expression of protein disulfide isomerase and heavy chain binding protein on antibody secretion in a recombinant CHO cell line. *Biotechnology progress*, 21(1), pp.106–11.

Bull, A.T., 2010. The renaissance of continuous culture in the post-genomics age. *Journal of industrial microbiology & biotechnology*, 37(10), pp.993–1021.

Choi, Y.S. et al., 2007. Enhancement of erythropoietin production in recombinant chinese hamster ovary cells by sodium lactate addition. *Biotechnology and Bioprocess Engineering*, 12(1), pp.60–72.

Chong, W.P.K. et al., 2010. Metabolomics-driven approach for the improvement of Chinese hamster ovary cell growth: overexpression of malate dehydrogenase II. *Journal of biotechnology*, 147(2), pp.116–21.

Chung, J.Y. et al., 2004. Effect of Doxycycline-Regulated Calnexin and Calreticulin Expression on Specific Thrombopoietin Productivity of Recombinant Chinese Hamster Ovary Cells. *Biotechnology and Bioengineering*, 85(5), pp.57.

Chusainow, J. et al., 2009. A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? *Biotechnology and Bioengineering*, 102(4), pp.1182–96.

Dietmair, S. et al., 2010. Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol. *Analytical Biochemistry*, 404(2), pp.155–164.

Hammond, S. et al., 2012. Chinese Hamster Genome Database: An Online Resource for the CHO Community at www.CHOgenome.org. *Biotechnology and bioengineering*, 109, pp.1353–1356.

Hansen, B.G. et al., 2011. Versatile enzyme expression and characterization system for Aspergillus nidulans, with the Penicillium brevicompactum polyketide synthase gene

from the mycophenolic acid gene cluster as a test case. *Applied and environmental microbiology*, 77(9), pp.3044–51.

Hayter, P.M. et al., 1992. {G}lucose-limited chemostat culture of chinese hamster ovary cells producing recombinant human interferon-gamma. *Biotechnology and Bioengineering*, 39(3), pp.327–335.

Hayter, P.M. et al., 1993. The effect of the dilution rate on CHO cell physiology and recombinant interferon-gamma production in glucose-limited chemostat culture. *Biotechnology and Bioengineering*, 42(9), pp.1077–1085.

Heyland, J. et al., 2010. Quantitative physiology of Pichia pastoris during glucose-limited high-cell density fed-batch cultivation for recombinant protein production. *Biotechnology and Bioengineering*, 107(2), pp.357–68.

Hung, F. et al., 2010. mRNA stability and antibody production in CHO cells: improvement through gene optimization. *Biotechnology journal*, 5(4), pp.393–401.

Hussain, H., Maldonado-Agurto, R. & Dickson, A.J., 2014. The endoplasmic reticulum and unfolded protein response in the control of mammalian recombinant protein production. *Biotechnology letters*, 36(8), pp.1581–93.

Hwang, S.O., Chung, J.Y. & Lee, G.M., 2003. Effect of doxycycline-regulated ERp57 expression on specific thrombopoietin productivity of recombinant CHO cells. *Biotechnology progress*, 19(1), pp.179–84.

Jayapal, K. & Wlaschin, K., 2007. Recombinant protein therapeutics from CHO cells-20 years and counting. *Chemical Engineering Progress*, 103(10), pp.40–47.

Jiang, Z., Huang, Y. & Sharfstein, S.T., 2006. Regulation of recombinant monoclonal antibody production in chinese hamster ovary cells: a comparative study of gene copy number, mRNA level, and protein expression. *Biotechnology progress*, 22(1), pp.313–8.

Jossé, L., Smales, C.M. & Tuite, M.F., 2012. Engineering the chaperone network of CHO cells for optimal recombinant protein production and authenticity. A. Lorence, ed. *Methods in molecular biology (Clifton, N.J.)*, 824, pp.595–608.

Kanehisa, M. et al., 2014. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research*, 42(Database issue), pp.D199–205.

Kanehisa, M. & Goto, S., 2000. KEGG : Kyoto Encyclopedia of Genes and Genomes. *Nucleic acids research*, 28(1), pp.27–30.

Kang, S. et al., 2014. Cell line profiling to improve monoclonal antibody production. *Biotechnology and Bioengineering*, 111(4), pp.748–60.

Kazemi Seresht, A. et al., 2013. Long-term adaptation of Saccharomyces cerevisiae to the burden of recombinant insulin production. *Biotechnology and Bioengineering*, 110(10), pp.2749–63.

Khoo, S.H.G., Falciani, F. & Al-Rubeai, M., 2007. A genome-wide transcriptional analysis of producer and non-producer NS0 myeloma cell lines. *Biotechnology and applied biochemistry*, 47(Pt 2), pp.85–95.

Kildegaard, H.F. et al., 2013. The emerging CHO systems biology era: harnessing the 'omics revolution for biotechnology. *Current opinion in biotechnology*, 24(6), pp.1102–7.

Kim, J. et al., 2004. Enhancement of Erythropoietin Production from Chinese Hamster Ovary (CHO) Cells by Introduction of the Urea Cycle Enzymes, Carbamoyl Phosphate Synthetase I and Ornithine Transcarbamylase. *J. microbiol. Biotechnol.*, 14(4), pp.844–851.

Kim, J.Y., Kim, Y.-G. & Lee, G.M., 2012. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Applied microbiology and biotechnology*, 93(3), pp.917–30.

Kim, Y.-G. et al., 2011. Effect of Bcl-xL overexpression on erythropoietin production in recombinant Chinese hamster ovary cells treated with dimethyl sulfoxide. *Process Biochemistry*, 46(11), pp.2201–2204.

Kim, Y.G. & Lee, G.M., 2009. Bcl-xL overexpression does not enhance specific erythropoietin productivity of recombinant CHO cells grown at 33 degrees C and 37 degrees C. *Biotechnology Progress*, 25(1), pp.252–256.

Lattenmayer, C. et al., 2007. Characterisation of recombinant CHO cell lines by investigation of protein productivities and genetic parameters. *Journal of biotechnology*, 128(4), pp.716–25.

Lattenmayer, C. & Loeschel, M., 2007. Protein-free transfection of CHO host cells with an IgG-fusion protein: Selection and characterization of stable high producers and comparison to conventionally transfected clones. *Biotechnology and Bioengineering*, 96(6), pp.1118–1126.

Lee, C.J. et al., 2009. A clone screening method using mRNA levels to determine specific productivity and product quality for monoclonal antibodies. *Biotechnology and Bioengineering*, 102(4), pp.1107–18.

Lee, F.W. et al., 1998. Engineering Chinese hamster ovary (CHO) cells to achieve an inverse growth - associated production of a foreign protein, beta-galactosidase. *Cytotechnology*, 28(1-3), pp.73–80.

Lee, Y.Y. et al., 2009. Overexpression of heat shock proteins (HSPs) in CHO cells for extended culture viability and improved recombinant protein production. *Journal of Biotechnology*, 143, pp.34–43.

Lund, A.M. et al., 2014. A versatile system for USER cloning-based assembly of expression vectors for mammalian cell engineering. *PloS one*, 9(5), p.e96693.

Mason, M. et al., 2012. Identifying bottlenecks in transient and stable production of recombinant monoclonal-antibody sequence variants in Chinese hamster ovary cells. *Biotechnology progress*, 28(3), pp.846–55.

Mead, E.J. et al., 2009. Identification of the limitations on recombinant gene expression in CHO cell lines with varying luciferase production rates. *Biotechnology and Bioengineering*, 102(6), pp.1593–602.

Mohan, C. & Lee, G.M., 2010. Effect of inducible co-overexpression of protein disulfide isomerase and endoplasmic reticulum oxidoreductase on the specific antibody productivity of recombinant Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 107(2), pp.337–46.

Niklas, J. et al., 2013. Metabolism and metabolic burden by alpha1-antitrypsin production in human AGE1.HN cells. *Metabolic Engineering*, 16(1), pp.103–114.

Nyberg, G.B. et al., 1999. Metabolism of peptide amino acids by Chinese hamster ovary cells grown in a complex medium. *Biotechnology and Bioengineering*, 62(3), pp.324–35.

O'Callaghan, P.M. et al., 2010. Cell line-specific control of recombinant monoclonal antibody production by CHO cells. *Biotechnology and Bioengineering*, 106(6), pp.938–51.

Peng, R.-W., Abellan, E. & Fussenegger, M., 2011. Differential effect of exocytic SNAREs on the production of recombinant proteins in mammalian cells. *Biotechnology and Bioengineering*, 108(3), pp.611–20.

Peng, R.-W. & Fussenegger, M., 2009. Molecular engineering of exocytic vesicle traffic enhances the productivity of Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 102(4), pp.1170–1181.

Powell, J. & Berkner, K., 1986. Human erythropoietin gene: high level expression in stably transfected mammalian cells and chromosome localization. *Proceedings of the Nation al Academy of Sciences of the United States of America*, 83, pp.6465–6469.

Pybus, L.P. et al., 2013. Model-directed engineering of "difficult-to-express" monoclonal antibody production by Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 111(2), pp.372–385.

Regenberg, B. et al., 2006. Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in Saccharomyces cerevisiae. *Genome Biology*, 7(11).

Reisinger, H. et al., 2008. The absence of effect of gene copy number and mRNA level on the amount of mAb secretion from mammalian cells. *Applied microbiology and biotechnology*, 81(4), pp.701–10.

Sellick, C. a. et al., 2010. Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: Matching extraction approaches to cell type and metabolite targets. *Metabolomics*, 6(3), pp.427–438.

Sengupta, N., Rose, S.T. & Morgan, J. a, 2011. Metabolic flux analysis of CHO cell metabolism in the late non-growth phase. *Biotechnology and Bioengineering*, 108(1), pp.82–92.

Smales, C.M. et al., 2004. Comparative proteomic analysis of GS-NS0 murine myeloma cell lines with varying recombinant monoclonal antibody production rate. *Biotechnology and Bioengineering*, 88(4), pp.474–88.

Sung, K.Y., Jong, K.H. & Gyun, M.L., 2004. Effect of simultaneous application of stressful culture conditions on specific productivity and heterogeneity of erythropoietin in Chinese hamster ovary cells. *Biotechnology Progress*, 20(4), pp.1293–1296.

Surabattula, R., Rao, K.R.S.S. & Polavarapu, R., 2011. An Optimized Process for Expression , Scale-Up and Purification of Recombinant Erythropoietin Produced in Chinese Hamster Ovary Cell Culture. *Research in Biotechnology*, 2(3), pp.58–74.

Untergasser, A. et al., 2012. Primer3--new capabilities and interfaces. *Nucleic acids research*, 40(15), p.e115.

Walsh, G., 2014. Biopharmaceutical benchmarks 2014. *Nature Biotechnology*, 32(10), pp.992–1000.

Yee, J.C., Gerdtzen, Z.P. & Hu, W.-S., 2009. Comparative transcriptome analysis to unveil genes affecting recombinant protein productivity in mammalian cells. *Biotechnology and Bioengineering*, 102(1), pp.246–263.

Yoon, S.K. et al., 2005. Effect of culture pH on erythropoietin production by Chinese hamster ovary cells grown in suspension at 32.5 and 37.0 degrees C. *Biotechnology and Bioengineering*, 89(3), pp.345–56.

Yoon, S.K., Song, J.Y. & Lee, G.M., 2003. Effect of low culture temperature on specific productivity, transcription level, and heterogeneity of erythropoietin in Chinese hamster ovary cells. *Biotechnology and Bioengineering*, 82(3), pp.289–98.

Zhou, H. et al., 2010. Generation of stable cell lines by site-specific integration of transgenes into engineered Chinese hamster ovary strains using an FLP-FRT system. *Journal of Biotechnology*, 147(2), pp.122–129.

TABLES

Table I. List of primers and corresponding sequences used for quantitation of gene copy numbers and gene expression levels.

Primer name	***Target gene***	***Purpose***	***Primer sequence 5'-3'***
EPO-Fwd	*hEPO*	Copy number determination	AGAGGCCGAGAACATCACCA
EPO-Rev	*hEPO*	Copy number determination	CCCACTTCCATCCGCTTA
GAPDH-Fwd	*Gapdh*	Copy number determination	AGCTTGTCATCAACGGGAAG
GAPDH-Rev	*Gapdh*	Copy number determination	ATCACCCCATTTGATGTT
ActB-Fwd	*Actb*	Copy number determination	CCAGCACCATGAAGATCAAG
ActB-Rev	*Actb*	Copy number determination	TGCTTGCTGATCCACATCTC
EPO (CHO optimized)-Fwd	*hEPO*	Plasmid construction	AGTGCGAUATGGGCGTGCACGAG TGTC
EPO (CHO optimized)-Rev	*hEPO*	Plasmid construction	AGACTGTGUTAATCTATCGCCGGT CCGGC

Table II. Raw data obtained in duplicate batch cultivations of EPO producing clones (C1-C7) and the parental clone (Control) in bioreactors. Abbreviations: **μmax** = maximum specific growth rate, **IVCD** = Integral of viable cell density (biomass yield), $\mathbf{q_{Glc}}$ = maximum specific glucose uptake rate, $\mathbf{q_{Lac}}$ = maximum specific lactate secretion rate, $\mathbf{q_{GLN}}$ = maximum specific glutamine uptake rate, $\mathbf{q_{GLU}}$ = maximum specific glutamate secretion rate, $\mathbf{q_{NH4}}$ = maximum specific ammonium secretion rate, $\mathbf{Y_{Lac/Glc}}$ = yield of lactate on glucose, $\mathbf{Y_{NH4/GLN}}$ = yield of ammonium on glutamine, $\mathbf{q_{EPO}}$ = specific EPO productivity, **Glc/GLN consumption** = uptake ratio of glucose per glutamine.

Clone	μ_{max} *[day^{-1}]*	***IVCD*** *[10^6 cells*h/mL]*	q_{Glc} *[pmol/cell/day]*	q_{Lac} *[pmol/cell/day]*	q_{Gln} *[pmol/cell/day]*	q_{Glu} *[pmol/cell/day]*	q_{NH4} *[pmol/cell/day]*	$Y_{Lac/Glc}$ *[mol/mol]*	$Y_{NH4/Gln}$ *[mol/mol]*	q_{EPO} *[pg/cell/day]*
C1	0.97 / 1.00	712 / 639	5.88 / 4.52	7.08 / 6.91	1.10 / 1.07	0.20 / 0.17	0.94 / 0.73	1.20 / 1.53	0.86 / 0.69	0.18 / 0.17
C2	0.91 / 1.05	627 / 587	6.73/ 5.28	7.14 / 7.59	0.96 / 1.05	0.16 / 0.16	0.80 / 0.64	1.06 / 1.44	0.84 / 0.61	0.36 / 0.21
C3	0.70 / 0.81	446 / 456	5.78 / 5.57	7.88 / 9.20	0.92 / 0.97	0.21 / 0.30	1.01 / 0.90	1.36 / 1.65	1.09 / 0.92	0.60 / 0.36

C4	0.86 / 0.97	544 / 532	4.58 / 4.69	6.88 / 7.17	1.00 / 1.06	0.23 / 0.18	0.88 / 0.78	1.50 / 1.53	0.88 / 0.73	0.76 / 0.54
C5	0.80 / 0.81	308 / 254	6.56 / 7.30	10.7/ 10.9	1.63 / 1.50	0.33 / 0.28	1.50 / 1.16	1.63 / 1.50	0.92 / 0.77	2.64 / 1.06
C6	0.89 / 1.00	585 / 579	5.86 / 5.18	5.96 / 7.42	1.05 / 1.00	0.20 / 0.21	0.87 / 0.72	1.02 / 1.43	0.82 / 0.71	3.05 / 3.39
C7	0.89 / 0.97	584 / 500	4.62 / 6.07	7.74 / 8.80	1.17 / 1.22	0.29 / 0.24	0.87 / 0.88	1.68 / 1.45	0.83 / 0.72	4.66 / 5.42
Control	**0.92 / 0.98**	**528 / 557**	**4.67 / 4.94**	**7.01 / 7.72**	**1.18 / 1.15**	**0.26 / 0.17**	**0.92 / 0.75**	**1.50 / 1.56**	**0.78 / 0.65**	-

Clone	*Glc/Gln Consumption* *[mol/mol]*	*Cell size* *[µm]*	*Aggregation rate* *[%]*
C1	5.36/ 4.26	13.41 / 13.30	29.52 / 15.85
C2	7.05 / 5.03	13.38 / 13.41	18.83 / 12.12
C3	6.27 / 5.73	13.05 / 12.90	25.85 / 11.97
C4	4.58 / 4.40	13.40 / 13.06	23.70 / 13.38
C5	4.02 / 4.87	16.25 / 16.01	35.04 / 32.18
C6	5.56 / 5.13	13.51 / 13.37	19.27 / 12.86
C7	3.95 / 4.96	13.94 / 13.69	20.03 / 8.61
Control	**3.95 / 4.30**	**13.92 / 13.37**	**22.64 / 11.81**

LIST OF FIGURES + LEGENDS

Figure 1. Specific EPO productivity. The error bars indicate standard deviation of two biological replicates**.**

Figure 2. Overview of intracellular energy and redox-related metabolites in EPO clones. A. Intracellular concentration of adenosine phosphates. **B.** Adenylate energy charge. **C.** Intracellular concentration of phosphorylated and non-phosphorylated nicotinamide adenine dinucleotides. **D.** Catabolic reduction charge. **E.** Anabolic reduction charge. Error bars indicate standard deviation of three biological replicates.

Figure 3. Schematic representation of glycolysis and associated levels of intracellular metabolites. Quantified metabolites are indicated with black font on the pathway map (left). The concentrations of 3-phosphoglycerate and 2-phosphoglycerate

were pooled, as they could not be separated in the method. Error bars indicate standard deviation of three biological replicates.

Figure 4. Culture dynamics of clone 1, clone 4 and clone 7 during 31 days of continuous culture in chemostat. A Viable cell densities. **B** Extracellular EPO titres. **C** Determined EPO gene copy numbers. **D** Determined EPO gene expression. **E** Ratio of determined EPO mRNA transcript per EPO gene (curves) and averaged specific EPO productivity for phase I, phase II and phase III (bars). **F** Ratio of secreted EPO per mRNA transcript (curves) and averaged specific EPO productivity for phase I, phase II and phase III (bars). Error bars indicate standard deviation of two biological replicates.

Figure 5. Differential gene expression analysis of amino acid catabolic genes in the high and low producer. A. Gene expression landscape of genes catalyzing the degradation or synthesis of amino acids. Circles indicate genes next to the reaction the encoded enzyme catalyzes. Gene expression values are shown as log fold-change indicating up- or down regulated genes in clone 7 relative to clone 1. Amino acids are colored blue, redox active metabolites are colored red and metabolites from the central metabolism are colored yellow. Reactions that do not produce or consume amino acids have been left out for simplicity. Dashed lines indicate multiple catalytic reactions. **B.** Frequency distribution of amino acids in human EPO without signal peptide. Black bars correspond to amino acids, which are preserved in clone 7 relative to clone 1. Grey bars indicate amino acids, which are not preserved. Red bars indicate amino acids that are secreted from the cells and therefore not considered in the analysis.

FIGURES

Figure 1

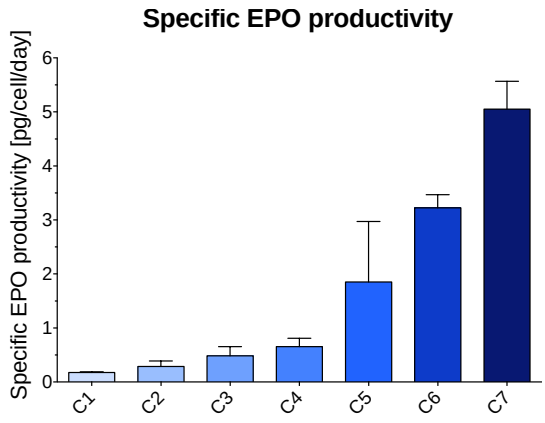


Figure 2

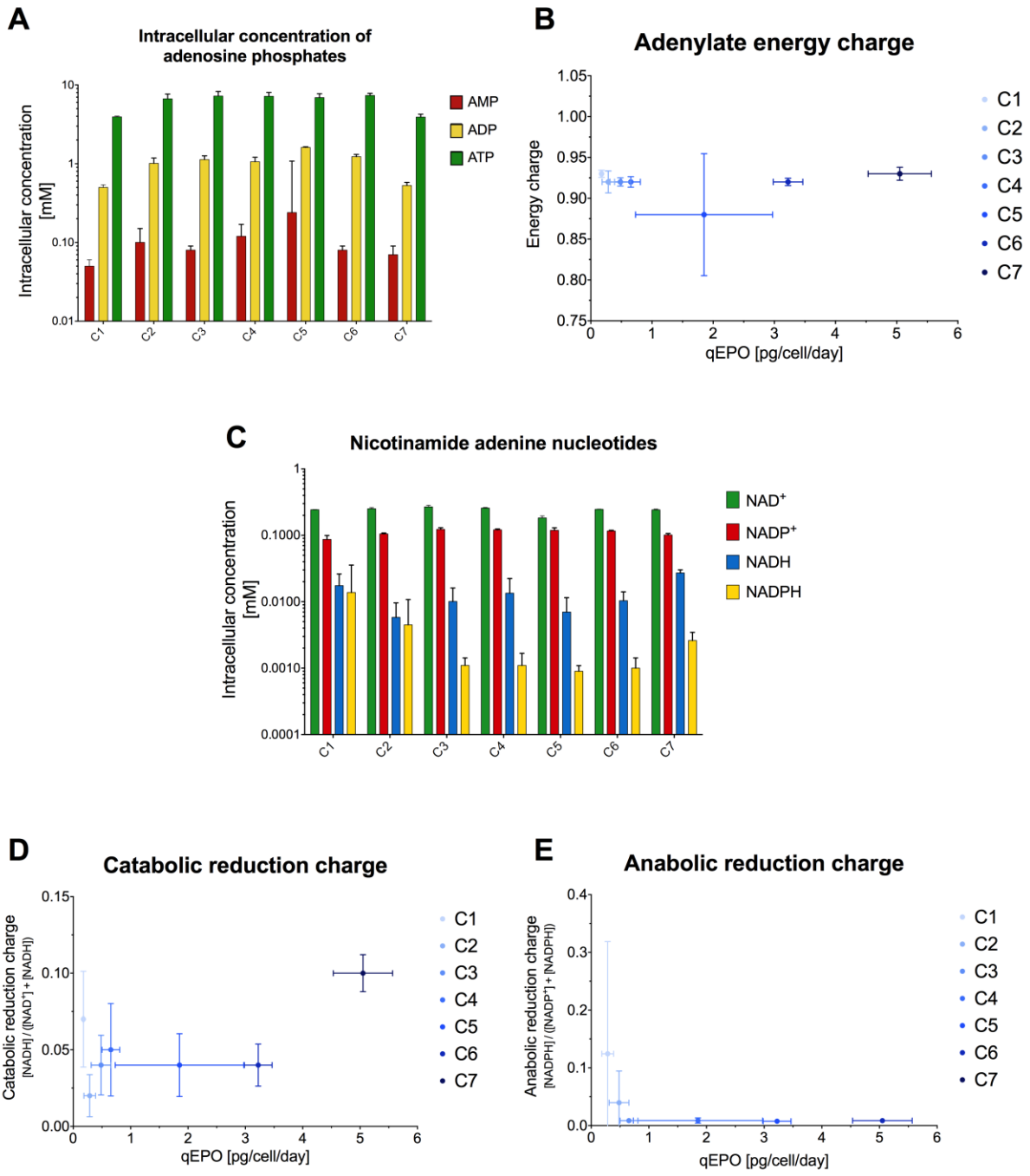

A
Intracellular concentration of adenosine phosphates
Intracellular concentration [mM]
AMP
ADP
ATP
C1
C2
C3
C4
C5
C6
C7
B
Adenylate energy charge
Energy charge
qEPO [pg/cell/day]
C
Nicotinamide adenine nucleotides
Intracellular concentration [mM]
NAD+
NADP+
NADH
NADPH
D
Catabolic reduction charge
Catabolic reduction charge [NADH] / ([NAD+] + [NADH])
qEPO [pg/cell/day]
E
Anabolic reduction charge
Anabolic reduction charge [NADPH] / ([NADP+] + [NADPH])
qEPO [pg/cell/day]

Figure 3

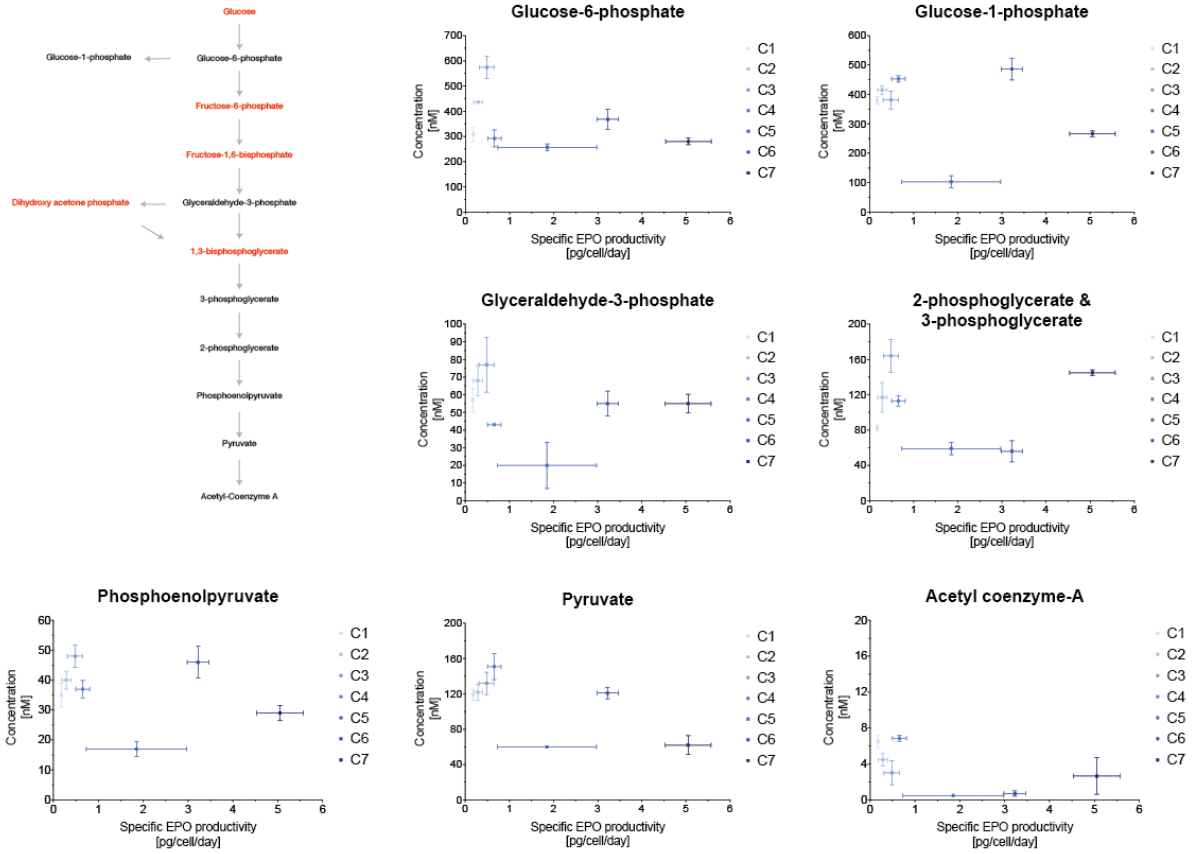
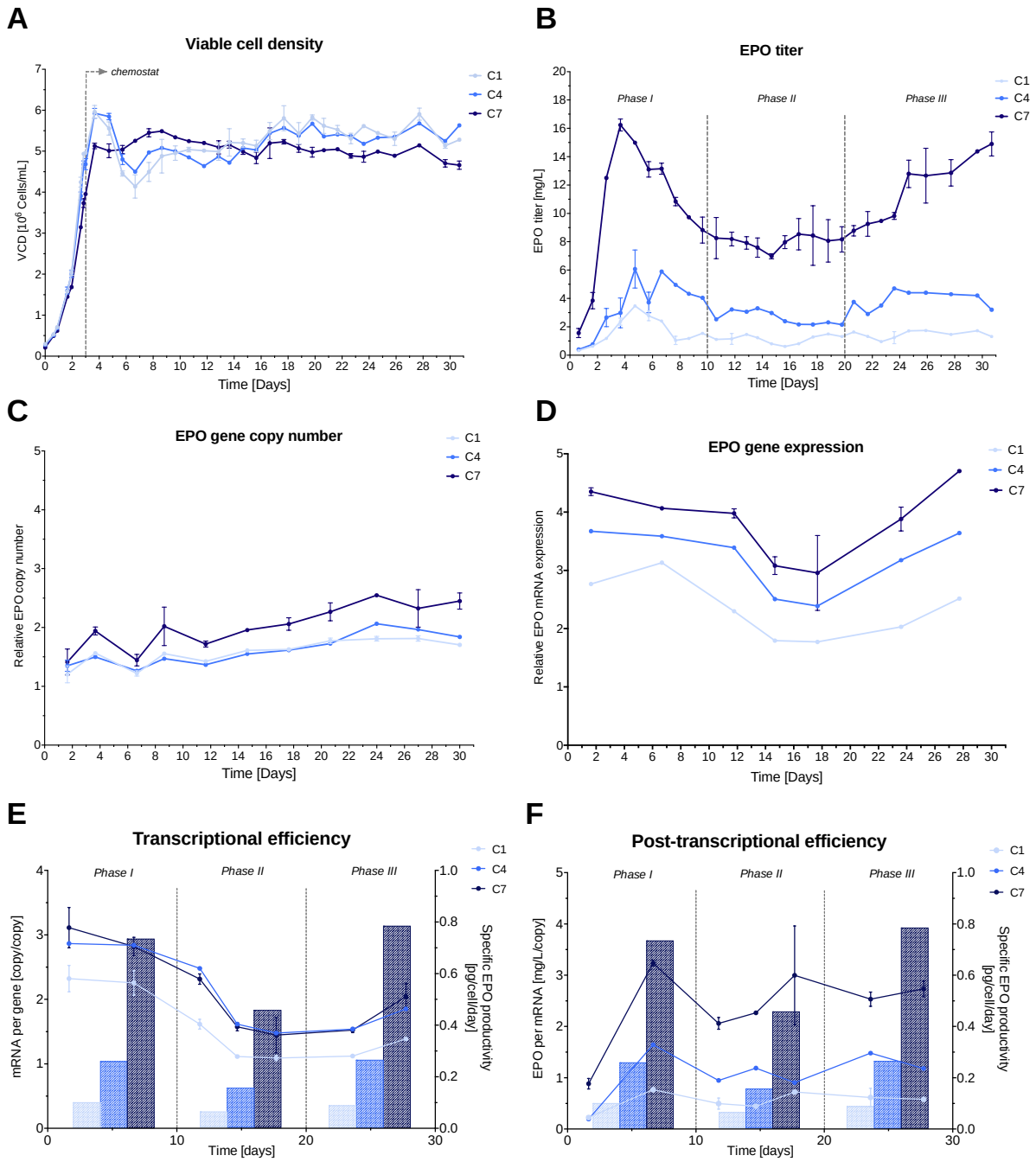
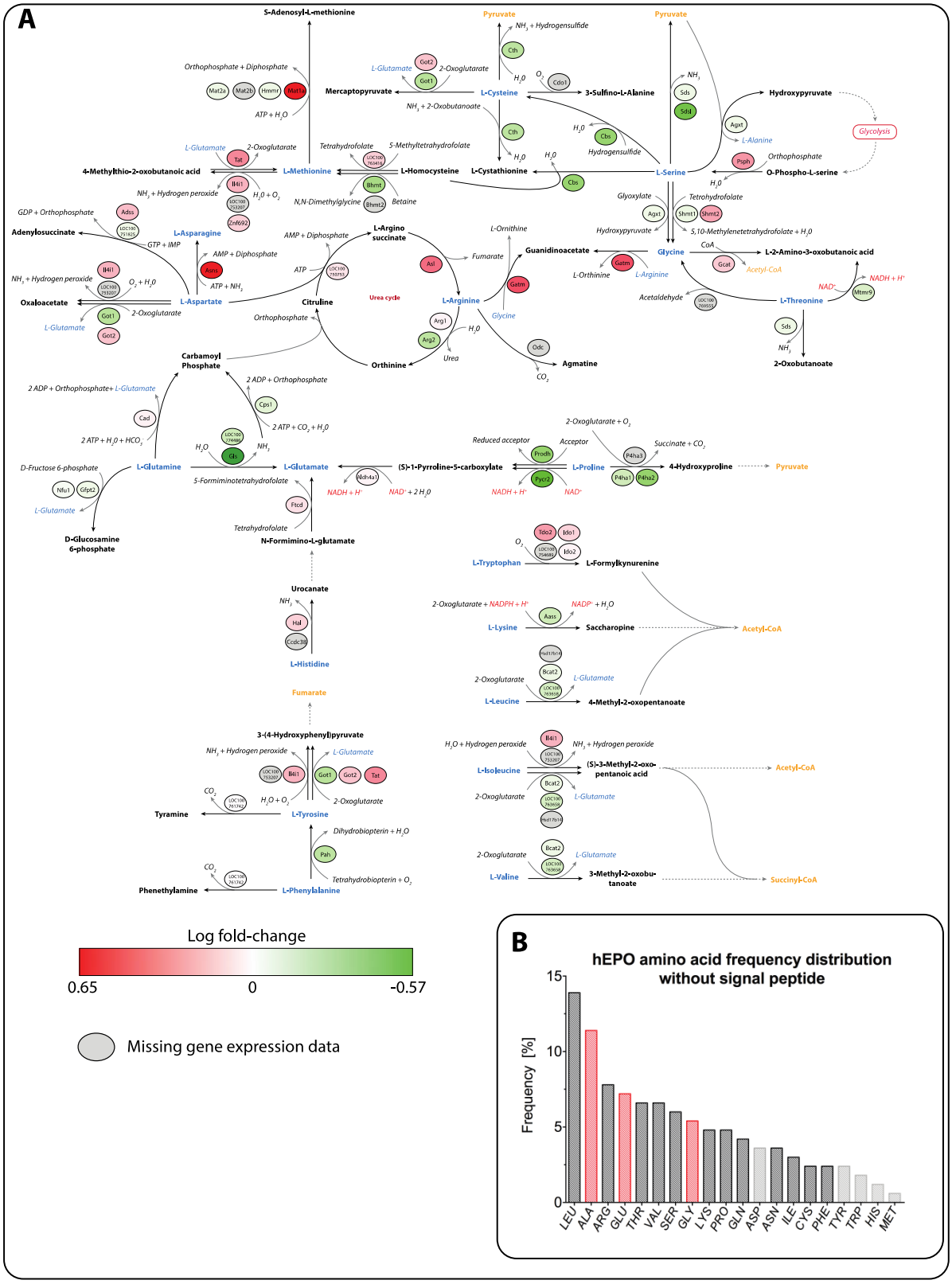

Glucose
Glucose-1-phosphate
Glucose-6-phosphate
Fructose-6-phosphate
Fructose-1,6-bisphosphate
Dihydroxy acetone phosphate
Glyceraldehyde-3-phosphate
1,3-bisphosphoglycerate
3-phosphoglycerate
2-phosphoglycerate
Phosphoenolpyruvate
Pyruvate
Acetyl-Coenzyme A
Glucose-6-phosphate
Glucose-1-phosphate
Glyceraldehyde-3-phosphate
2-phosphoglycerate & 3-phosphoglycerate
Phosphoenolpyruvate
Pyruvate
Acetyl coenzyme-A
Concentration [nM]
Specific EPO productivity [pg/cell/day]
C1
C2
C3
C4
C5
C6
C7

Figure 4


A
Viable cell density
chemostat
C1
C4
C7
VCD [10^6 Cells/mL]
Time [Days]
B
EPO titer
Phase I
Phase II
Phase III
EPO titer [mg/L]
Time [Days]
C
EPO gene copy number
Relative EPO copy number
Time [Days]
D
EPO gene expression
Relative EPO mRNA expression
Time [Days]
E
Transcriptional efficiency
mRNA per gene [copy/copy]
Specific EPO productivity [pg/cell/day]
Time [days]
F
Post-transcriptional efficiency
EPO per mRNA [mg/L/copy]
Specific EPO productivity [pg/cell/day]
Time [days]

1 **Figure 5**


A
S-Adenosyl-L-methionine
Pyruvate
Pyruvate
Mercaptopyruvate
L-Cysteine
3-Sulfino-L-Alanine
Hydroxypyruvate
Glycolysis
4-Methylthio-2-oxobutanoic acid
L-Methionine
L-Homocysteine
L-Cystathionine
L-Serine
O-Phospho-L-serine
Adenylosuccinate
L-Asparagine
L-Argino succinate
Guanidinoacetate
Glycine
L-2-Amino-3-oxobutanoic acid
Oxaloacetate
L-Aspartate
Citruline
Urea cycle
L-Arginine
L-Ornithine
L-Threonine
2-Oxobutanoate
Carbamoyl Phosphate
Orthinine
Agmatine
L-Glutamine
L-Glutamate
(S)-1-Pyrroline-5-carboxylate
L-Proline
4-Hydroxyproline
Pyruvate
D-Glucosamine 6-phosphate
N-Formimino-L-glutamate
Urocanate
L-Histidine
L-Tryptophan
L-Formylkynurenine
L-Lysine
Saccharopine
Acetyl-CoA
L-Leucine
4-Methyl-2-oxopentanoate
Fumarate
3-(4-Hydroxyphenyl)pyruvate
L-Tyrosine
Tyramine
L-Phenylalanine
Phenethylamine
(S)-3-Methyl-2-oxo-pentanoic acid
L-Isoleucine
L-Valine
3-Methyl-2-oxobu-tanoate
Succinyl-CoA
Log fold-change
0.65
0
-0.57
Missing gene expression data
B
hEPO amino acid frequency distribution without signal peptide
Frequency [%]
15
10
5
0
LEU
ALA
ARG
GLU
THR
VAL
SER
GLY
LYS
PRO
GLN
ASP
ASN
ILE
CYS
PHE
TYR
TRP
HIS
MET

2

6.4 Paper 4 - Benchmarking two commonly used *Saccharomyces cerevisiae* strains for heterologous vanillin-β-glucoside production

Tomas Strucko, **Olivera Magdenoska**, Uffe Mortensen**.**

Paper submitted to Metabolic Engineering Communications 2015

Benchmarking two commonly used *Saccharomyces cerevisiae* strains for heterologous vanillin-β-glucoside production

Authors

Tomas Strucko[1], Olivera Magdenoska[1] and Uffe H Mortensen[1]*

Affiliations

1 – Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

*Corresponding author

Phone: +45 45 25 27 01

FAX: + 45 45 88 41 48

E-mail: um@bio.dtu.dk

Abstract

The yeast *Saccharomyces cerevisiae* is a widely used eukaryotic model organism and a key cell factory for production of biofuels and wide range of chemicals. From the broad palette of available yeast strains, the most popular are those derived from laboratory strain S288c and the industrially relevant CEN.PK strain series. Importantly, in recent years these two strains have been subjected to comparative "-omics" analyses pointing out significant genotypic and phenotypic differences. It is therefore possible that the two strains differ significantly with respect to their potential as cell factories for production of specific compounds. To examine this possibility, we have reconstructed a *de novo* vanillin-β-glucoside pathway in an identical manner in S288c and CEN.PK strains. Characterization of the two resulting strains in two standard conditions revealed that the S288c background strain produced up to 10-fold higher amounts of vanillin-β-glucoside compared to CEN.PK. This study demonstrates that yeast strain background may play a major role in the outcome of newly developed cell factories for production of a given product.

***Keywords*:** Yeast; Cell factory; Strain choice; Heterologous production; Vanillin-glucoside; Shikimate pathway.

1 Introduction

The recent achievements in the field of systems biology and metabolic engineering combined with a wide array of molecular biology tools has established the yeast *Saccharomyces cerevisiae* as a key cell factory for heterologous production of scientifically and industrially relevant products. The latter comprise a large variety of products ranging from low-value bulk chemicals and biofuels (e.g., ethanol) to food additives (e.g., flavors and colorants) and high value pharmaceuticals (e.g., recombinant proteins) [1], [2]. Today, a range of different *S. cerevisiae* strain backgrounds are available for the yeast community of which BY (S288c), W303 and CEN.PK are the most frequently used [3]. The variety of strains has been developed by different laboratories to suit a range of diverse research goals within different disciplines such as genetics, physiology and biochemistry. For example, CEN.PK strain [4], is a popular platform for physiological as well as metabolic engineering studies whereas S288c, the first eukaryote to be sequenced [5], was mainly used for genetic studies, but has in recent years been increasingly used as an alternative platform for metabolic engineering experiments [1], [2]. Specifically, recent surveys show that over the past ten years the two strains were used in more than 50% of the analyses with CEN.PK series being the most popular (approx. 37% - CEN.PK *vs.* 24% - S288c and its derivatives) [1], [2]. The importance of the two strains as cell factories is further substantiated by an extensive multi-laboratory efforts which were made for systematic comparisons of S288c and CEN.PK [6]–[8]. Genetic differences of these strains were revealed first by microarray studies [9] and more recently genotype to phenotype relation was investigated after whole-genome sequencing of the CEN.PK113-7D strain [10], [11]. For example, the sequencing comparison studies revealed that 83 genes were absent in CEN.PK relative to S288C [11] and that more than 22,000 single nucleotide polymorphisms (SNPs) exist between the two strains. 13,000 of the SNPs are distributed in 1,843 open reading frames (ORFs) and the activity of a large number of proteins may therefore differ in the two strains; especially since 35 % of these SNPs result in amino-acid residue substitutions. The remaining 9,000 SNPs, which are mainly distributed in the intergenic regions, may potentially impact gene expression and thereby protein levels in the two strains. The fact that the highest enrichment of SNPs was detected in genes involved in carboxylic acid, organic acid, and carbohydrate metabolism, as well as, in nitrogen, amino acid, lipid and aromatic compound metabolism suggests that the basic metabolism in the two strain background may be quite different [10]. Importantly, based on the combined "-omics" analyses, several phenotypic differences between the two strains were assigned to these mutations, e.g., differences in galactose uptake and ergosterol biosynthesis, etc. in two strains [10].

The fact that the two main *S. cerevisiae* backgrounds for construction of cell factories are genetically/phenotypically quite different raises the possibility that heterologous production of a

given compound in CEN.PK and in S288c may result in significantly different yields. To examine this possibility, we therefore constructed comparable cell factories for vanillin-β-glucoside (VG) production in S288c and CEN.PK backgrounds. Importantly, all five relevant genes for VG production in *S. cerevisiae*, (see **Figure 1** and Hansen et al.[12]) were integrated at phenotypically neutral and well defined locations in the yeast chromosome using a recently published integration platform [13]. Moreover, in order for the engineered strains to be directly comparable, auxotrophies were eliminated by a sexual backcross to the corresponding wild-type strains. Remarkably, the physiological characterizations of both VG cell factories in two different cultivation modes revealed major differences in the VG production. Next, we examined the comprehensive "-omics" datasets for S288c and CEN.PK strains [7] to look for plausible reasons for the different VG production profiles in the two genetic backgrounds. Together our analyses serve as a step towards a scenario where the optimal genetic background for cell factory construction can be successfully selected based on a systems biology model for yeast cell factories.

2 Materials and Methods

2.1 DNA cloning procedures

The DNA fragments used for vector construction were amplified by PCR with PfuX7 polymerase developed by Nørholm et al. [14] using the primers listed in the **Table 2**. Molecular coning was done by uracil-specific excision reagent (USER™) as previously described in [15], [16].Final constructs were validated by sequencing (StarSEQ® GmbH, Germany). The genes constituting the *de novo* VG pathway (1-*3DSD*, 2- *ACAR*, 3-*EntD*, 4-*HsOMT*, and 5-*UGT*) and a set of bidirectional promoter (pPGK1/pTEF1) were amplified by PCR from the appropriate vector templates (see **Table 1**). The pathway genes and promoters were assembled into vectors designed to integrate on chromosome XII [13]. A total of three plasmids were produced: pXII1-23 (pPGK1::ACAR; pTEF1::EntD), pXII2-54 (pPGK1::UGT; pTEF1::HsOMT) and pXII5-01 (pTEF1::3DSD), see **Figure S1**. For a full list of plasmids used or constructed in this work see **Table 1**.

2.2 Strain construction

The genotype and source of the strains used in this study is given in the **Table 3**. Two different background strains CEN.PK113-11C and X2180-1A (isogenic to S288c [17]) were used as hosts for reconstruction of the de novo vanillin-β-glucoside (VG) pathway. All yeast constructs were generated by high efficiency transformation method (lithium acetate/polyethylene glycol/single carrier DNA) previously described by Gietz et al. [18].

To construct two yeast strains containing the VG pathway, vectors pXII1-23, pXII2-54 and pXII5-01 harboring the following genes (*ACAR* and *EntD*), (*UGT* and *HsOMT*) and (*3DSD*), respectively, were

digested with *Not*I restriction enzyme (Fermentas-Thermo Fisher Scientific) and gel-purified using illustra GFX PCR DNA and Gel Band Purification Kit (GE Lifesciences). The individual gene targeting substrates were transformed iteratively into both yeast strain backgrounds in three consecutive transformations. To prevent undesired production of toxic intermediates especially protocatechuic aldehyde (PAL), the *3DSD* gene was cloned at the latest step. After each round of transformation the URA3 marker was eliminated by direct repeat recombination and counter-selection on 5-FOA [19] allowing the URA3 marker to be recycled.

Complete gene deletions of *BGL1* and *ADH6* were achieved using method described by Güldener et al [20]. PCR fragment carrying *loxP*-KanMX-*loxP* flanked by 40 nt long segment homologous to sequences of Up- and Down-stream of the appropriate open reading frame (ORF) to be deleted were amplified form plasmid pUG6. The KanMX marker was excised by expressing Cre recombinase from the vector pSH47 [20]. All gene targeting events were validated by diagnostic PCR using specially designed primer pairs (see **Table 2**).

To eliminate any auxotrophies, two engineered strains S-VG-aux and C-VG-aux were sexually crossed to S288c and CEN.PK110-16D, respectively. This resulted in two final prototrophic vanillin-β-glucoside producing yeast strains S-VG (S288c based) and C-VG (CEN.PK based). Genetic cross and selection procedures were performed as described in [21]. Schematic flowchart representing the strain construction is depicted in **Figure S2**.

2.3 Media

For cloning purposes lysogeny broth (LB) [22] supplemented with 100 mg/L of ampicillin (Sigma-Aldrich) was used for growing of *Escherichia coli* DH5α.

All media used for genetic manipulations of yeast were prepared as previously described by Sherman et al. [23], with minor modifications of synthetic medium where leucine concentration was doubled to 60mg/L. All yeast transformants with gene integrations were selected on synthetic complete media missing uracil (SC-ura). For subsequent round of transformations URA3 marker was recycled through direct repeat recombination and selected on synthetic complete media containing 30 mg/L uracil and 740mg/L 5-fluoroorotic acid (5-FOA) (Sigma-Aldrich).

Yeast transformants with necessary gene deletions were selected on Yeast Extract Peptone Dextrose (YPD) plates supplemented with 200 mg/L of G418 (Sigma-Aldrich) [20]. The medium composition is as follows: 10 g/L of yeast extract, 20 g/L of peptone and 20 g/L of glucose and 20 g/L of agar.

A defined minimal medium previously described by Verduyn [24] with glucose as a carbon source was used for *S. cerevisiae* cultivations in batch and chemostat cultures. For batch cultivations the

medium was supplemented with 30 g/L of glucose [7], whereas for feed media for continuous cultivations contained 7.5 g/L of glucose. The medium was composed of: 7.5 g/L $(NH_4)_2SO_4$, 3 g/L KH_2PO_4, 0.75 g/L Mg_2SO_4, 1.5 mL/L trace metal solution, 1.5 mL/L vitamins solution, 0.05 mL/L antifoam 204 (Sigma-Aldrich). Trace metal solution contains 3 g/L $FeSO_4 \cdot 7H_2O$, 4.5 g/L $ZnSO_4 \cdot 7H_2O$, 4.5 g/L $CaCl_2 \cdot 6H_2O$, 0.84 g/L $MnCl_2 \cdot 2H2O$, 0.3 g/L $CoCl_2 \cdot 6H_2O$, 0.3 g/L $CuSO_4 \cdot 5H_2O$, 0.4 g/L $NaMoO_4 \cdot 2H_2O$, 1 g/L H_3BO_3, 0.1 g/L KI and 15 g/L $Na_2EDTA \cdot 2H_2O$. Vitamins solution includes 50 mg/L d-biotin, 200 mg/L para-amino benzoic acid, 1.0 g/L nicotinic acid, 1.0 g/L Ca-pantothenate, 1.0 g/L pyridoxine HCL, 1.0 g/L thiamine HCl and 25 mg/L minositol. Glucose was autoclaved separately and vitamins solutions were sterile filtered (pore size 0.2 µm Ministart®-Plus, Sartorius AG, Germany) and added after autoclavation.

2.4 Batch and chemostat cultivations

For each biological replica separate colonies of engineered yeast strains from YPD plate were inoculated to the 0.5 L shake flasks with 50 mL of the previously described minimal medium (pH 6.5). Pre-cultures were incubated in an orbital shaker set to 150 rpm at 30°C until mid-exponential phase $OD_{600} \approx 5$ and directly used for inoculation. In this experiment, batch cultivations were performed in duplicates and continuous cultures – in triplicates.

Batch cultivations were performed under aerobic conditions in 1L fermenters equipped with continuous data acquisition (Sartorius, B. Braun Biotech International, GmbH, Melsungen, Germany) with a working volume of 1L. Fermenters were inoculated with initial $O.D._{600}$ =0.05. To ensure aeration a stirrer speed was set to 600 rpm and airflow rate to 1.0 v.v.m. (60 L/h). The temperature was maintained at 30°C during the cultivation and pH=5.0 level was controlled by automatic addition of 2M NaOH or 2M H_2SO_4. The batch cultures were sampled in regular intervals through both glucose and ethanol growth phases. Glucose-limited chemostat cultures were grown in the same conditions as previously described for batch cultivations. Chemostats were initiated as batch cultivations with starting glucose concentration of 15 g/L and OD_{600} =0.05 and switched to continuous mode in early exponential phase. Minimal medium was fed at a constant dilution rate of 0.1 h^{-1} [7].The working volume of 1 L was kept constant by a level based outlet. Samples were taken after at least five residence times (50 hours) of constant biomass and carbon dioxide concentration readings.

In both cultivation experiments exhaust gas composition was constantly monitored by off gas analyzer (1311 Fast response triple gas, Innova) combined with Mass Spectrometer (Prima Pro Process MS, Thermo Fisher Scientific, Winsford UK). No elevated ethanol concentration in the exhaust (less than 1.5% of Ethanol evaporation) was detected.

The biomass concentration was monitored by measuring both optical density at 600nm wavelength (OD_{600}) and cell dry weight (DW) in the cultivation broth. OD_{600} was estimated using a UV mini 1240 spectrophotometer (Shimadzu, Kyoto, Japan), biomass samples were diluted with distilled water to achieve OD_{600} reading within 0.1 to 0.4 range. DW measurements were performed using polyethersulfone (PES) filters with a pore size of 0.45 µm Montamil® (Membrane Solutions, LLC). The filters were pre-dried in a microwave oven at 150 W and weighed. A known volume of cultivation broth (5 mL) was filtered and then washed with approx. 15 mL of distilled water. Finally, the filters with biomass were dried in the microwave oven at 150 W and cell DW was determined [25].

2.5 Extracellular metabolite measurements

Extracellular metabolites were determined by high performance liquid chromatography (HPLC) analysis. Two distinct HPLC methods were applied for analysis of different groups of extracellular metabolites.

The submerged cultivation samples for yeast primary metabolites were centrifuged at 12000xg for 2 min, supernatant was transferred to a new tube and stored at -20 ^{o}C until further analysis. Glucose, ethanol, glycerol, pyruvate, succinate and acetate were determined by high performance liquid chromatography (HPLC) analysis using an Aminex HPX-87H ion-exclusion column (Bio-Rad Laboratories, Hercules, CA). The column temperature was kept at 60 ^{o}C and the elution was performed using 5 mM H_2SO_4 with constant flow rate of 0.6 mL min^{-1}. Metabolite detection was performed by a RI-101differential refractometer detector (Shodex) and an UVD340U absorbance detector (Dionex) set at 210 nm.

Samples for quantification of vanillin-β-glucoside and its pathway metabolites were prepared as follows: 1mL of cultivation broth and 1mL of 96% EtOH was carefully mixed by vortex and centrifuged at 12000xg for 2 min, supernatant was transferred to a new tube and stored at -20^{o}C until further analysis. Extracellular vanillin-β-glucoside (VG), vanillin (VAN), protocatechuic acid (PAC), protocatechuic aldehyde (PAL) and vanillic acid (VAC) were quantified using Agilent 1100 series equipment with a Synergi Polar-RP 150*2 mm 4u column (Phenomenex). A gradient of acetonitrile (ACN) with 1% tetra-fluoroacetic acid (TFA) and water with 1% TFA at a constant flow rate of 0.5 mL/min was used as mobile phase. The elution profile was as follows: 5% ACN — 1 min, 5% ACN to 30% ACN — 8 min, 30% ACN to 100% ACN — 1 min, 100% ACN — 1 minute, 100% ACN to 5% ACN — 3 min. The column was kept at 40 ^{o}C and metabolite detection was performed using a UV diode-array detector set to 230 and 280 nm.

2.6 Intracellular metabolite measurements

Samples for intracellular co-factor metabolites were taken during steady state conditions at the end of continuous cultivation. The samples were quenched and extracted as previously described by Villas-Boas et al. [26]; 5 mL of culture broth was sprayed into pre-cooled (-40°C) falcon tube containing 20 mL of 60% methanol, spun down for 2 min at 5000xg in precooled centrifuge (-10°C) and extracted using boiling ethanol method [26] followed by evaporation under nitrogen. The samples were re-dissolved in 300 µL eluent A (10 mM tributylamine and 10 mM acetic acid). The analysis was carried out on an Agilent 1290 binary UHPLC system coupled with an Agilent 6460 triple quadrupole mass spectrometer (Torrance, CA, USA). The MS was operated in negative ion and multiple reactions monitoring mode. Separation of 0.5 µL samples was performed by ion-pair chromatography, as described in details in Magdenoska et al. [27] using 10 mM tributylamine as ion pair reagent. The gradient used was: 0-12 min 0-50% B, 12-12.5 min 50-100% B, 12.5-14 min 100% B, 14-14.5 min 100-0% B, 14.5-19.5 min 0% B. External standard calibration method was used for quantification. The calibration curves were constructed by preparing calibration solutions ranging from 1 to 100 µg/mL for ATP and 0.3 to 25 µg/mL for NADPH and UDP-glucose. Both the quenched extracts and the calibration solutions were spiked with 60 µL mixture containing 150 µg/mL [U-^{13}C] ATP. The quantification was carried out using Mass Hunter Quantitative analysis software (version B.06.00).

3 Results

3.1 *De novo* pathway reconstruction in CEN.PK and S288c backgrounds

To compare CEN.PK and S288c for their ability to produce VG in a fair manner, it was necessary to insert all genes identically in both strain backgrounds. In the original VG producing strain, the individual genes of the pathway were inserted by repeated integrations into the *TPI1* promoter region in an S288c based strain [12]. Consequently, the gene order is ill-defined and not easy to reconstruct in CEN.PK. We therefore introduced the VG pathway in the two strain backgrounds, S288c (isogenic isolate X2180-1A [17]) and CEN.PK113-11C [4] in a defined manner. Specifically, the five genes used by Hansen et al [12] for VG production were integrated into three sites located on chromosome XII (see **Figure 2**), which are part of a defined gene expression platform we have previously established [13]. Importantly, prior to integration, we compared the up- and downstream sequences at these integration sites in CEN.PK and S288c and found that they differ by only a single SNP, a C in CEN.PK and G in S288c, which is present in the upstream targeting sequence of the XII-5 site. The five genes are therefore inserted into a genetic context, which is essentially identical. Additionally, two genes *ADH6* and *EXG1* were deleted in both strains to minimize unwanted side reactions with VG pathway metabolites. Finally, for both genetic backgrounds, auxotrophic markers

were eliminated by sexual back crossing to wild-type variants of the two strain backgrounds to produce two prototrophic VG producing strains, C-VG (CEN.PK based) and S-VG (S288c based), which we used for further analysis.

3.2 Physiological characterization of vanillin-β-glucoside producing strains during batch and continuous fermentations

Before evaluating the VG production ability of the two strain backgrounds, we first assessed whether VG production affected the overall physiology of the C-VG and S-VG strains. Hence, they were grown in batch and as continuous cultures in well-controlled bioreactors under standard laboratory cultivation conditions [7]. Like for wild-type strains, the overall growth profiles exhibited by the S-VG and C-VG strains in batch reactors were composed by two growth phases. One initial growth phase where all glucose was fermented (GF – phase); followed, after the diauxic shift, by a second growth phase where all accumulated ethanol was respiro-fermented (ER – phase), see **Figure 3**. We note that the specific biomass yield on glucose (**Table 4**) appeared higher (14%) in C-VG as compared to S-VG, but in our experiment this difference was not significant ($p > 0.33$). A similar difference has previously been observed for the parental S288c and CEN.PK strains grown at the same conditions [3], [7]. It has previously been shown that CEN.PK grows faster than S288c at these conditions, $\mu_{max} \approx 0.4\ h^{-1}$ vs. $\mu_{max} \approx 0.3\ h^{-1}$, respectively [7], [10]. In contrast, C-VG and S-VG grew slower than the corresponding wild-type strains and both strains displayed identical growth rates on glucose ($\mu_{max} \approx 0.2\ h^{-1}$). Further analysis of the cultivation broth obtained at different time points during growth, showed that the production profiles of five primary metabolites (ethanol, pyruvate, succinate, glycerol, and acetate) in the central carbon metabolism of S-VG and C-VG were similar to what has previously been observed with the corresponding wild-type strains (**Figure S3**). Among the remaining parameters —r_{Glc}, r_{Eth}, and r_{Gly}, only the latter varied between the two strains as it was approx. three-fold higher in the S-VG ($p < 0.002$). Finally, we also note that S-VG displayed a growth deficiency during the ER phase which was nearly twice as long with S-VG (23 h) than with C-VG (12 h) partly due to a much longer delay from the diauxic shift to exponential growth is resumed. This deficiency has also been previously reported for wild-type S288c [10].

In chemostats, steady-state conditions with constant production of biomass as well as stable readings of carbon dioxide and oxygen by the off-gas analyzer where obtained for both strain backgrounds. For the C-VG strain this was achieved in less than 5 residence times (50 hours) after feeding was initiated, whereas for the S-VG strain it took more than 8 residence times (80 hours), see **Figure S4**. Importantly, at a dilution rate of $0.1\ h^{-1}$, both strains propagated exclusively by glucose respiratory metabolism [28] as no production of ethanol, glycerol and organic acids was observed (see **Table 4).** The C-VG strain produced significantly more (seven percent, $p < 0.01$)

biomass as compared to S-VG, on the contrary, the specific glucose uptake rate was approx. seven % higher ($p < 0.01$) in S-VG strain.

3.3 High VG yields with S288c are generated during its prolonged ethanol respiratory growth phase

VG production in the S-VG and C-VG strains was initially compared at 45 h after both cultures have reached stationary phase. In the original VG producing strain constructed by Hansen et al [12], various levels of intermediates were observed due to an unbalanced heterologous VG pathway [12], [29]. In agreement with this, analyses of the extracellular metabolite levels in the cultivation broth showed that PAC, PAL, VAC, and VAN accumulated in both S-VG and C-VG in batch cultivation. The yields of each of the intermediates were similar in the two strains (**Figure 4**) with PAC being by far the most prominent metabolite. In fact, PAC accounted for ~70- and 75% of the carbon ending up as intermediates in C-VG and S-VG strains, respectively. However, when the two strains were analyzed for VG production, we surprisingly observed that twice as much VG was produced with the S-VG strain than with the C-VG strain ($p < 0.05$). This finding prompted us to determine and compare the production of VG and its intermediates of the two strains during the different growth phases throughout the entire batch cultivation (**Figure 5**A-B). The VG yields were almost identical with the two strains during the GF phase (**Figure 5**C). The total carbon ending up in VG pathway metabolites (TCV) is 14% lower in C-VG as compared to S-VG. This effect is mainly due to less accumulated PAC in C-VG (37% less in C-VG, $p < 0.05$). In contrast, with S-VG, a two-fold higher ($p < 0.05$) VG yield was generated during the ER phase than with C-VG despite that the amounts of intermediates accumulating at the end of the phase were similar in both strains (see **Figure 5**C). As a result, TCV was 45% higher with S-VG as compared to C-VG. Yields and productivities on glucose during exponential growth and during steady state condition are represented in the Table S1).

To further understand why VG yields were higher with S-VG than with C-VG, we determined whether production of VG and VG-intermediates was proportional to biomass during all time-points. Based on this analysis, we observed for both strains that the efficiencies of TCV, VAC and VG production were higher in the ER phase as compared to the GF phase. In contrast, for both strains PAC was produced with equal efficiency in the two phases **Figure 6**. After the diauxic shift, C-VG quickly entered a new state where these metabolites were produced in amounts proportional to biomass. With S-VG such states are also achieved for TCV, PAC and VAC although, for TCV and VAC, more time was required for these states to be reached in this strain as compared to C-VG. In contrast, with S-VG a state where VG was produced proportionally to biomass was never achieved as the VG production efficiency increased during the entire phase, see **Figure 6**D.

3.4 Steady-state VG production at glucose limitation is higher in S288c than in CEN.PK

Next, we investigate VG production in continuous cultures where we could obtain strictly glucose respiratory conditions for both strains, see above. Dramatically, this analysis revealed that the amounts of TCV and VG were significantly increased to levels four- and ten-fold higher with S-VG than with the C-VG strain ($p < 8.8E-9$, $p < 1.1E-10$, respectively), see **Figure 7**. When the levels of intermediates were inspected, we observed that with S-VG, the majority (68%) of the TCV ended up in VG, whereas the rest of the carbon ended up in PAC (18%) and in VAC (13 %). In contrast, VG constituted only 20 % of the TCV in the C-VG strain. The remaining carbon was mainly ending up in VAC (approx. 70 %) and only little (less than 5 %) in PAC.

Four co-factors, ATP, NADPH, SAM and UDP-Glc, are used for formation of VG. In the paper by Canelas et al. ([7]) where "-omics" data for the two background strains were compared, the levels of three of these cofactors (ATP, NADPH and UDP-Glc) were measured. We were able to quantify ATP and therefore determined the concentrations of this metabolite in C-VG and S-VG to investigate whether ATP levels were altered due to the presence of the VG pathway. Analysis of samples obtained at steady state showed that concentrations of ATP were 10.6 µmol/g DW and 8.5 µmol/g DW for C-VG and S-VG, respectively. These numbers are somewhat higher than the corresponding numbers reported by Canelas et al. and this may be due to lab to lab differences. Importantly, the relative ATP levels of the two strains (C-VG to S-VG) is 1.2 in our experiment as well as in theirs [7]. This indicates that the presence of the VG pathway does not change ATP levels in the two strains despite that 10-fold more VG is formed in S-VG as compared in to C-VG. Due to the absence of internal standards for NADPH and UDP-Glc, we were not able to compare our data with those in the literature. However, UDP-Glc levels were 0.5 µmol/g DW for both strains. These results were also confirmed by a high resolution mass spectrometer, where similar chemical profiles were obtained for both strains. The NADPH levels measured in C-VG and S-VG in our study were close to the limit of detection preventing a comparison between the two strains.

4 Discussion

The importance of *Saccharomyces cerevisiae* in the development of novel cell factories is demonstrated by the large number of industrially relevant substances that can now be produced in this organism [1], [2], [31]. Several laboratory yeast strains have been used for this purpose with S288c and CEN.PK being the most popular [2], [7]. However, despite that the two strains are genetically and physiologically very different [3], [7], [10], [11]; these differences are rarely used to

determine, which strain background should be chosen as a cell factory for *de novo* production of a given compound. In this study, we therefore investigated whether choice of yeast strain background for production of vanillin-β-glucoside (VG) is an issue that can be advantageously considered. Our finding that heterologous production of VG was dramatically more efficient in an S288c based strain compared to CEN.PK in both batch and continuous cultivations demonstrates that this is indeed the case.

Our physiological characterizations of the C-VG and S-VG strains show that the S-VG strain produces significantly more VG in both batch and continuous cultivations. One explanation could be a higher flux of carbon into the VG pathway in the S-VG strain. In agreement with this, both the first intermediate in the VG pathway (PAC) and TCV accumulate to significant higher levels with S-VG as compared to with C-VG during both continuous and batch cultivations, see **Figure 4** and **Figure 7**. For batch cultivation, we note this is true at all time points examined, see **Figure 6**.

Carbon for VG production is recruited from the shikimate pathway, which is a part of the aromatic amino acid biosynthesis and, which is well characterized in yeast [32], [33]. The pathway is tightly regulated via two 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase isoenzymes Aro3 and Aro4, which are feedback inhibited by phenylalanine and tyrosine, respectively (**Figure 1**). In this context, we note that in the case of CEN.PK derived strains, the flux through the shikimate pathway is increased more than four-fold if feedback inhibition is eliminated [34]; and that this feature has been successfully exploited in a metabolic strategy to increase the yield of CEN.PK strains producing the plant flavonoid naringenin, which is derived from phenylalanine and tyrosine [35]. Interestingly, two of the genes in the shikimate pathway, *ARO1* and *ARO3*, contain ten SNPs that result in amino acid substitutions. Moreover, a large number of SNPs are distributed in the up- and downstream regions of the genes in the shikimate pathways. It is therefore likely that the level and/or activities of enzymes of the shikimate pathway are different in the two strain backgrounds. In agreement with this, the available "-omics" datasets presented by Canelas [7] show that the intracellular concentration of shikimic acid (the direct metabolite of 3-DHS) is significantly higher in S288c than in CEN.PK strains at similar conditions in batch and continuous cultivations. Moreover, they also reported that under batch conditions CEN.PK strains are able to maintain higher intracellular amino-acid pools, which was suggested, and later shown, to be a result of increased protein turnover rate [36].

In VG producing strains, we therefore speculate that higher concentrations of intracellular aromatic amino acids in C-VG might inhibit Aro3 and Aro4 leading to a decreased flux towards 3-DHS and consequently to a lower capacity for VG production. When the genomes of S288c and CEN.PK SNPs

were compared, several SNPs exist in genes encoding for enzymes involved in the amino acid and aromatic compound metabolism. For example, we have compared ORFs sequences of several *ARO* genes in the strains, see **Table 5**. Importantly, the gene of the pentafunctional enzyme Aro1 (which is directly involved in synthesis of 3-DHS) contains 15 SNPs, of which, seven results in non-synonymous amino acid substitutions. Moreover, the gene of the regulatory protein Aro3 was found to contain 23 SNPs, where three caused amino acid substitutions. It is therefore possible that the flux through the shikimate pathway could be different in the two strain backgrounds.

Other features in the two strain backgrounds could influence VG production. For example, in the case of batch cultivation, the higher VG yields in S-VG strains are mainly created after the diauxic shift. This suggests that the S-VG strain has increased activity of the aromatic biosynthetic pathway only at respiratory metabolism, which is consistent with the even higher activity in the glucose limited chemostat cultures. This could be due to a limitation in the supply of erythrose-4-phosphate, a precursor for aromatic amino acid biosynthesis and an intermediate of the pentose-phosphate (PP) pathway. In this context, we note that at fermentative conditions the flux through the PP pathway is much lower than at respiratory conditions [37].

One feature that could influence the final VG yield is availability of co-factors that are required to convert PAC to VG. To this end we note that ATP and UDP-glucose levels seem to be unaffected by the VG pathway in the two strain backgrounds. Moreover, we have previously introduced mutations designed to increase the NADPH/NADP level in the original VG producing strain (S288c background) and this resulted in higher VG levels showing that cofactor availability appears limiting. However, we note that in this study, the organization of the inserted VG pathway genes, including gene copy numbers, in the producer strain is not known [29]. We did not measure SAM - levels in C-VG and S-VG and this co-factor may vary between the two strains. Similarly, the activity of the PPTase, EntD, which activates ACAR by covalently attaching phosphopantheine to the apoenzyme is unknown and may differ in the two strain backgrounds. Considering the more than 13,000 SNPs between CEN.PK and S288c, one may also expect that differences in the global metabolism may influence VG production. For example, 20 genes in the mitogen-activated protein kinase (MAPK) signaling pathway contain SNPs resulting in amino acid residue substitutions. Among those, one is in Cyr1, which changes the activity through the pathway controlled by the global regulator protein kinase A, PKA [11], [38], [39]. The ability of the two strains to adapt to toxic intermediates, especially PAL and VAN [12], may also vary and influence yields. Even if they do not, accumulation of these intermediates vary slightly in the two strain back grounds and may therefore differentially influence

VG yields. To this end, we note that the growth rates for C-VG and S-VG are reduced as compared to the corresponding strains that do not contain the VG pathway.

Since, the VG yield for S-VG is approximately 20-fold below the maximum theoretical yield; the effects described above are unlikely due to carbon being channeled into the VG pathway at the expense of other destinations and functions. An interesting question that still remains to be answered is therefore how the two strain backgrounds develop as VG producers as yields are improved by e.g. metabolic engineering. Similarly, it would be interesting to address whether some of the obvious genetic differences pointed out above can be transferred from S-VG to C-VG to improve the VG yields in the latter.

In conclusion, we have shown that heterologous production of VG differs dramatically in two different strain backgrounds. Moreover, our analyses taking advantage of the "-omics" data presented in the literature suggest that it may be possible to predict, which of the two strain backgrounds that would be the better producer. As additional data constantly accumulate we expect that such qualified guesses can be made in an increasingly safe manner, but with the present insights, we recommend to test more than one genetic background during construction of a novel cell factory.

Acknowledgements

We thank Jens Nielsen and Morten Kielland-Brandt for comments on this manuscript, Kristian Fog Nielsen for supporting metabolite analyses; and Anna-Lena Heins and Ted Johansen for facilitating fermentation experiments. This work was supported by grant 0603-00323B from The Danish Council for Strategic Research.

References

[1] I.-K. Kim, A. Roldão, V. Siewers, and J. Nielsen, "A systems-level approach for metabolic engineering of yeast cell factories.," *FEMS Yeast Res.*, vol. 12, no. 2, pp. 228–48, Mar. 2012.

[2] K.-K. Hong and J. Nielsen, "Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries.," *Cell. Mol. Life Sci.*, vol. 69, no. 16, pp. 2671–90, Aug. 2012.

[3] van Dijken JP, J. Bauer, L. Brambilla, P. Duboc, J. Francois, C. Gancedo, M. Giuseppin, J. Heijnen, M. Hoare, H. Lange, E. Madden, P. Niederberger, J. Nielsen, J. Parrou, T. Petit, D.

Porro, M. Reuss, van Riel N, M. Rizzi, H. Steensma, C. Verrips, J. Vindeløv, and J. Pronk, “An interlaboratory comparison of physiological and genetic properties of four Saccharomyces cerevisiae strains.,” *Enzyme Microb. Technol.*, vol. 26, no. 9–10, pp. 706–714, Jun. 2000.

[4] K. Entian and P. Kötter, “Yeast genetic strain and plasmid collections,” *Methods Microbiol.*, vol. 36, no. 06, pp. 629–666, 2007.

[5] A. Goffeau, B. Barrell, H. Bussey, and R. Davis, “Life with 6000 genes,” *Science (80-.).*, vol. 274, no. 5287, 1996.

[6] a Rogowska-Wrzesinska, P. M. Larsen, A. Blomberg, A. Görg, P. Roepstorff, J. Norbeck, and S. J. Fey, “Comparison of the proteomes of three yeast wild type strains: CEN.PK2, FY1679 and W303.,” *Comp. Funct. Genomics*, vol. 2, no. 4, pp. 207–25, Jan. 2001.

[7] A. B. Canelas, N. Harrison, A. Fazio, J. Zhang, J.-P. Pitkänen, J. van den Brink, B. M. Bakker, L. Bogner, J. Bouwman, J. I. Castrillo, A. Cankorur, P. Chumnanpuen, P. Daran-Lapujade, D. Dikicioglu, K. van Eunen, J. C. Ewald, J. J. Heijnen, B. Kirdar, I. Mattila, F. I. C. Mensonides, A. Niebel, M. Penttilä, J. T. Pronk, M. Reuss, L. Salusjärvi, U. Sauer, D. Sherman, M. Siemann-Herzberg, H. Westerhoff, J. de Winde, D. Petranovic, S. G. Oliver, C. T. Workman, N. Zamboni, and J. Nielsen, “Integrated multilaboratory systems biology reveals differences in protein metabolism between two reference yeast strains.,” *Nat. Commun.*, vol. 1, no. 9, p. 145, Jan. 2010.

[8] M. T. a P. Kresnowati, W. a van Winden, M. J. H. Almering, a ten Pierick, C. Ras, T. a Knijnenburg, P. Daran-Lapujade, J. T. Pronk, J. J. Heijnen, and J. M. Daran, “When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation.,” *Mol. Syst. Biol.*, vol. 2, p. 49, Jan. 2006.

[9] P. Daran-Lapujade, J. Daran, P. Kotter, T. Petit, M. Piper, and J. Pronk, “Comparative genotyping of the laboratory strains S288C and CEN.PK113-7D using oligonucleotide microarrays,” *FEMS Yeast Res.*, vol. 4, no. 3, pp. 259–269, Dec. 2003.

[10] J. M. Otero, W. Vongsangnak, M. a Asadollahi, R. Olivares-Hernandes, J. Maury, L. Farinelli, L. Barlocher, M. Osterås, M. Schalk, A. Clark, and J. Nielsen, “Whole genome sequencing of Saccharomyces cerevisiae: from genotype to phenotype for improved metabolic engineering applications.,” *BMC Genomics*, vol. 11, no. 1, p. 723, Jan. 2010.

[11] J. F. Nijkamp, M. van den Broek, E. Datema, S. de Kok, L. Bosman, M. a Luttik, P. Daran-Lapujade, W. Vongsangnak, J. Nielsen, W. H. M. Heijne, P. Klaassen, C. J. Paddon, D. Platt, P. Kötter, R. C. van Ham, M. J. T. Reinders, J. T. Pronk, D. de Ridder, and J.-M. Daran, "De novo sequencing, assembly and analysis of the genome of the laboratory strain Saccharomyces cerevisiae CEN.PK113-7D, a model for modern industrial biotechnology.," *Microb. Cell Fact.*, vol. 11, no. 1, p. 36, Jan. 2012.

[12] E. E. H. Hansen, B. B. L. Møller, G. R. G. Kock, C. M. Bünner, C. Kristensen, O. R. Jensen, F. T. Okkels, C. E. Olsen, M. S. Motawia, and J. Hansen, "De novo biosynthesis of vanillin in fission yeast (Schizosaccharomyces pombe) and baker's yeast (Saccharomyces cerevisiae).," *Appl. Environ. Microbiol.*, vol. 75, no. 9, pp. 2765–74, May 2009.

[13] M. D. Mikkelsen, L. D. Buron, B. Salomonsen, C. E. Olsen, B. G. Hansen, U. H. Mortensen, and B. A. Halkier, "Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform.," *Metab. Eng.*, vol. 14, no. 2, pp. 104–11, Mar. 2012.

[14] M. H. H. Nørholm, "A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering.," *BMC Biotechnol.*, vol. 10, p. 21, Jan. 2010.

[15] F. Geu-Flores, H. H. Nour-Eldin, M. T. Nielsen, and B. a Halkier, "USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products.," *Nucleic Acids Res.*, vol. 35, no. 7, p. e55, Jan. 2007.

[16] H. H. Nour-Eldin, B. G. Hansen, M. H. H. Nørholm, J. K. Jensen, and B. a Halkier, "Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments.," *Nucleic Acids Res.*, vol. 34, no. 18, p. e122, Jan. 2006.

[17] R. K. Mortimer and J. R. Johnston, "Genealogy of principal strains of the yeast genetic stock center.," *Genetics*, vol. 113, no. 1, pp. 35–43, May 1986.

[18] D. Gietz, A. St Jean, R. a Woods, and R. H. Schiestl, "Improved method for high efficiency transformation of intact yeast cells.," *Nucleic Acids Res.*, vol. 20, no. 6, p. 1425, Mar. 1992.

[19] J. D. Boeke, F. Lacroute, and G. R. Fink, "Short communication A positive selection for mutants lacking orotidine-5 ' -phosphate decarboxylase activity in yeast : 5-fluoro-orotic acid resistance," *Mol. Gen. Genet.*, pp. 345–346, 1984.

[20] U. Güldener, S. Heck, T. Fielder, J. Beinhauer, and J. H. Hegemann, “A new efficient gene disruption cassette for repeated use in budding yeast.,” *Nucleic Acids Res.*, vol. 24, no. 13, pp. 2519–24, Jul. 1996.

[21] D. a Treco and F. Winston, “Growth and manipulation of yeast.,” *Curr. Protoc. Mol. Biol.*, vol. Chapter 13, no. April, p. Unit 13.2, Apr. 2008.

[22] G. Bertani, “Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli.,” *J. Bacteriol.*, vol. 62, no. 3, pp. 293–300, Sep. 1951.

[23] F. Sherman, G. R. Fink, and J. B. Hicks, *Laboratory course manual for methods in yeast genetics*. Cold Spring Harbor Laboratory, 1986.

[24] C. Verduyn, E. Postma, W. a Scheffers, and J. P. Van Dijken, “Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation.,” *Yeast*, vol. 8, no. 7, pp. 501–17, Jul. 1992.

[25] L. Olsson and J. Nielsen, “On-line and in situ monitoring of biomass in submerged cultivations,” *Trends Biotechnol.*, vol. 15, no. December, pp. 517–522, 1997.

[26] S. G. Villas-Bôas, J. Højer-Pedersen, M. Akesson, J. Smedsgaard, and J. Nielsen, “Global metabolite analysis of yeast: evaluation of sample preparation methods.,” *Yeast*, vol. 22, no. 14, pp. 1155–69, Oct. 2005.

[27] O. Magdenoska, J. Martinussen, J. Thykaer, and K. F. Nielsen, “Dispersive solid phase extraction combined with ion-pair ultra high-performance liquid chromatography tandem mass spectrometry for quantification of nucleotides in Lactococcus lactis.,” *Anal. Biochem.*, vol. 440, no. 2, pp. 166–77, Sep. 2013.

[28] P. Van Hoek, “Effect of specific growth rate on fermentative capacity of baker’s yeast,” *Appl. Environ. …*, vol. 64, no. 11, pp. 4226–4233, 1998.

[29] A. R. Brochado, C. Matos, B. L. Møller, J. Hansen, U. H. Mortensen, and K. R. Patil, “Improved vanillin production in baker’s yeast through in silico design.,” *Microb. Cell Fact.*, vol. 9, no. 1, p. 84, Jan. 2010.

[30] H. C. Lange, M. Eman, G. van Zuijlen, D. Visser, J. C. van Dam, J. Frank, M. J. de Mattos, and J. J. Heijnen, “Improved rapid sampling for in vivo kinetics of intracellular metabolites in Saccharomyces cerevisiae.,” *Biotechnol. Bioeng.*, vol. 75, no. 4, pp. 406–15, Nov. 2001.

[31] E. Nevoigt, “Progress in metabolic engineering of Saccharomyces cerevisiae.,” *Microbiol. Mol. Biol. Rev.*, vol. 72, no. 3, pp. 379–412, Sep. 2008.

[32] S. Teshiba, R. Furter, P. Niederberger, G. Braus, G. Paravicini, and R. Hütter, “Cloning of the ARO3 gene of Saccharomyces cerevisiae and its regulation.,” *Mol. Gen. Genet.*, vol. 205, no. 2, pp. 353–7, Nov. 1986.

[33] G. H. Braus, “Aromatic amino acid biosynthesis in the yeast Saccharomyces cerevisiae: a model system for the regulation of a eukaryotic biosynthetic pathway.,” *Microbiol. Rev.*, vol. 55, no. 3, pp. 349–70, Sep. 1991.

[34] M. a H. Luttik, Z. Vuralhan, E. Suir, G. H. Braus, J. T. Pronk, and J. M. Daran, “Alleviation of feedback inhibition in Saccharomyces cerevisiae aromatic amino acid biosynthesis: quantification of metabolic impact.,” *Metab. Eng.*, vol. 10, no. 3–4, pp. 141–53, 2008.

[35] F. Koopman, J. Beekwilder, B. Crimi, A. van Houwelingen, R. D. Hall, D. Bosch, A. J. a van Maris, J. T. Pronk, and J.-M. Daran, “De novo production of the flavonoid naringenin in engineered Saccharomyces cerevisiae.,” *Microb. Cell Fact.*, vol. 11, p. 155, Jan. 2012.

[36] K.-K. Hong, J. Hou, S. Shoaie, J. Nielsen, and S. Bordel, “Dynamic 13C-labeling experiments prove important differences in protein turnover rate between two Saccharomyces cerevisiae strains.,” *FEMS Yeast Res.*, vol. 12, no. 7, pp. 741–7, Nov. 2012.

[37] O. Frick and C. Wittmann, “Characterization of the metabolic shift between oxidative and fermentative growth in Saccharomyces cerevisiae by comparative ^{13}C flux analysis.,” *Microb. Cell Fact.*, vol. 4, p. 30, Nov. 2005.

[38] M. Vanhalewyn, Ë. Dumortier, S. Colombo, P. Ma, P. Van Dijck, and M. Johan, “Saccharomyces cerevisiae adenylate cyclase, Cyr1K1876M, specifically affects glucose-and acidification-induced cAMP signalling and not the basal cAMP,” *Mol. …*, vol. 33, pp. 363–376, 1999.

[39] K. Tanaka, M. Nakafuku, T. Satoh, M. S. Marshall, J. B. Gibbs, K. Matsumoto, Y. Kaziro, and a Toh-e, “S. cerevisiae genes IRA1 and IRA2 encode proteins that may be functionally equivalent to mammalian ras GTPase activating protein.,” *Cell*, vol. 60, no. 5, pp. 803–7, Mar. 1990.

[40] S. Partow, V. Siewers, S. Bjørn, J. Nielsen, and J. Maury, “Characterization of different promoters for designing a new expression vector in Saccharomyces cerevisiae,” *Yeast*, no. July, pp. 955–964, 2010.

[41] S. Nakagawa, Y. Niimura, T. Gojobori, H. Tanaka, and K. Miura, “Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes.,” *Nucleic Acids Res.*, vol. 36, no. 3, pp. 861–71, Feb. 2008.

[42] D. R. Cavener and S. C. Ray, “Eukaryotic start and stop translation sites.,” *Nucleic Acids Res.*, vol. 19, no. 12, pp. 3185–92, Jun. 1991.

[43] H. C. Lange and J. J. Heijnen, “Statistical reconciliation of the elemental and molecular biomass composition of Saccharomyces cerevisiae.,” *Biotechnol. Bioeng.*, vol. 75, no. 3, pp. 334–44, Nov. 2001.

Table 1. List of the plasmids used and constructed in this study.

Name	Genetic element	Reference
pXII1-23	*pPGK1::ACAR, pTEF1::EntD*	This study
pXII2-54	*pPGK1::UGT, pTEF1::HsOMT*	This study
pXII5-01	*pTEF1::3DSD*	This study
pXII1	—	Mikkelsen et al. [13]
pXII2	—	Mikkelsen et al. [13]
pXII5	—	Mikkelsen et al. [13]
pSP-G2	*pPGK1;pTEF1*	Partow et al. [40]
pJH500	*3DSD*	Hansen et al. [12]
pJH674	*ACAR*	Hansen et al. [12]
pJH589	*EntD*	Hansen et al. [12]
pJH543	*HsOMT*	Hansen et al. [12]
pJH665	*UGT*	Hansen et al. [12]
pUG6	*KanMX*	Güldener et al [20]
pSH47	*Cre recombinase*	Güldener et al [20]

Table 2. List of the primers used in this study. All sequences are presented in 5' to 3' direction, standard capital letter are gene specific sequences, **bold** letters represent USER specific tails, regular underline letters represent targeting sequences for appropriate gene deletions. *Italic* letters represent translational enhancer sequence [41], [42].

Name	Sequence
Primers for cloning purposes	
PGK_R-	**ACCCGTTGAU**GCCGCTTGTTTTATATTTGTTG
TEF_F+	**CGTGCGAU**GCCGCACACACCATAGCTTC
TEF_R+	**ACGTATCGCU**GTGAGTCGTATTACGGATCCTTG
DSD_F+	**AGCGATACGU***AAAA*ATGCCTTCCAAACTCGCC
DSD_R+	**CACGCGAU**TTACAAAGCCGCTGACAGC
ACAR_F-	**ATCAACGGGU***AAAA*ATGGCTGTTGATTCACCAGATG
ACAR_R-	**CGTGCGAU**CTTATAACAATTGTAACAATTCCAAATC
hOMT_F+	**AGCGATACGU***AAAA*ATGGGTGACACTAAGGAGCAA
hOMT_R+	**CACGCGAU**CTTATGGACCAGCTTCAGAACC
PPT1_F+	**AGCGATACGU***AAAA*ATGGTCGATATGAAAACTACGC
PPT1_R+	**CACGCGAU**TTAATCGTGTTGGCACAGC
UGT1_F-	**ATCAACGGGU***AAAA*ATGCATATCACAAAACCACACG
UGT1_R-	**CGTGCGAU**ACTAGGCACCACGTGACAAGTC
Deletion primers	
BGL1_del_F	ATTTTTTGTTTACTTTCTTTTTCTAGTTAATTACCAACTAAACTTCGTACGCTGCAGGTC
BGL1_del_R	CATTAGAAAATTCAGCTAAAATGAGCGGACTGAGGGCGACTAGTGGATCTGATATCACCTA
ADH6_del_F	GAGGAAGAAATTCAACACAACAACAAGAAAAGCCAAAATCCTTCGTACGCTGCAGGTC
ADH6_del_R	GTTAAAAAGAAAGGAGCTACATTTATCAAGAGCTTGACAACTAGTGGATCTGATATCACCTA
Verification primers	
XII-1-up-out-sq	CTGGCAAGAGAACCACCAAT
XII-2-up-out-sq	CGAAGAAGGCCTGCAATTC
XII-5-up-out-sq	CCACCGAAGTTGATTTGCTT
C1_TADH1_F	CTTGAGTAACTCTTTCCTGTAGGTC
BGL1_VF	TCATCCTCCCTGTGTTTACA
BGL1_VR	AGTTGAAACAGAGGATAAGGTG
ADH6_VF	GTTTTGCTTTTTTCTCTGGG
ADH6_VR	GGAGTATCAACCACTAAAGCG
Primers for sequencing	
S1_TEF_F	CGGTCTTCAATTTCTCAAGTTTC
S1_PGK_F	CAAGGGGGTGGTTTAGTTTAGT
S1_ACAR_F	CATTGTTCGTTCCAGACACTGAC
S2_ACAR_F	GAAATTGTCTCAAGGTGAGTTCG
S3_ACAR_F	TTTGGGTAGATTCTTGTGTTTGG
YF19	AAAAAATAAATAGGGACCTAGACTTCA
YF21	GACCTACAGGAAAGAGTTACTCAAGAAT

Table 3. List of the yeast strains used in this study.

Name	Genotype	Reference
X2180-1A (ura3)	*MATa SUC2 mal mel gal2 CUP1 ura3-52*	Public domain [17]
S288c	*MATα SUC2 gal2 mal mel flo1 flo8-1 hap1 ho bio1 bio6*	ATCC 204508
CEN.PK113-11C	*MATa MAL2-8C SUC2 ura3-52 his3Δ*	Peter Kötter[1]
CEN.PK110-16D	*MATα MAL2-8C SUC2 trp1-289*	Peter Kötter[1]
C-VG-aux	*MATa MAL2-8C SUC2 ura3-52 his3Δ XII2(pTEF1-HsOMT, pPGK1-UGT) XII1(pTEF1-PPT, pPGK1-ACAR) XII5(pTEF1-3DSD) Δbgl1::loxP Δadh6::KanMX*	This study
S-VG-aux	*MATa SUC2 gal2 mal mel ura3-52 XII2(pTEF1-HsOMT, pPGK1-UGT) XII1(pTEF1-PPT, pPGK1-ACAR) XII5(pTEF1-3DSD) Δbgl1::loxP Δadh6::KanMX*	This study
C-VG	*MATα MAL2-8C SUC2 XII2(pTEF1-HsOMT, pPGK1-UGT) XII1(pTEF1-PPT, pPGK1-ACAR) XII5(pTEF1-3DSD) Δbgl1::loxP Δadh6::KanMX*	This study
S-VG	*MATα SUC2 gal2 mal mel XII2(pTEF1-HsOMT, pPGK1-UGT) XII1(pTEF1-PPT, pPGK1-ACAR) XII5(pTEF1-3DSD) Δbgl1::loxP Δadh6::KanMX*	This study

[1] Institut für Mikrobiologie, der Johan Wolfgang Goethe-Universität, Frankfurt am Main, Germany.

Table 4. Physiologic parameters of the two engineered strains. Y_{sx} – biomass yield on glucose, r – specific metabolite production or consumption rates (C-mmol/g(DW)·h); Glc –glucose, Eth – ethanol and Gly – glycerol. NA – not applicable and ND – not detected. The yield coefficient Y_{sx} for biomass is calculated based on a molecular weight for biomass of 26.4 g/C-mol [43]. Errors represent standard deviation, 2≤n≤3.

Cultivation mode	Batch		Chemostat	
Strain	S-VG	C-VG	S-VG	C-VG
Glucose μ, h^{-1}	0.209 ± 0.002	0.199 ± 0.002	(0.1)*	(0.1)*
Ethanol μ, h^{-1}	0.05 ± 0.01	0.08 ± 0.01	NA	NA
Y_{sx} (C-mol/C-mol)	0.133 ± 0.014	0.151 ± 0.001	0.56 ± 0.02	0.60 ± 0.02
$-r_{Glc}$	60.7 ± 8.0	49.9 ± 0.6	6.81 ± 0.29	6.37 ± 0.20
r_{Eth}	25.9 ± 3.1	20.9 ± 3.0	ND	ND
r_{Gly}	4.05 ± 0.05	1.40 ± 0.13	ND	ND

* - the dilution rate used in this study.

Table 5. Point mutations of *ARO* genes found in CEN.PK113-7D compared to S228c. Mutations represented by one letter code for amino acid and number denoting the position in the protein.

Genes	SNPs		Amino acid substitution
	Total	Non-synonymous	
ARO1	15	7	T225I, P337S, S517P, N844T, M1141K, V1386I, G1576A
ARO3	23	3	K141R, E214D, S349T
ARO4	—	—	—
ARO7	—	—	—

Figure 1. *De novo* vanillin-β-glucoside pathway in *S. cerevisiae* assembled by Hansen and co-workers [12]. Colored boxes represent metabolites of the pathway: PAC – protocatechuic acid, PAL – protocatechuic aldehyde, VAC – vanillic acid, VAN – vanillin, VG – vanillin-glucoside. Black thick arrows show enzymatic reactions by heterologous enzymes: 3DSD – 3-dehydroshikimate dehydratase (*Podospora pauciseta*), ACAR – aromatic carboxylic acid reductase (*Neurospora sp*.), EntD – phosphopantetheine transferase (*Escherichia coli*), HsOMT – O-methyltransferase (*Homo sapiens*) and UGT – UDP-glycosyltransferase (*Arabidopsis thaliana*). Grey arrows and boxes (left side of the picture) represent simplified shikimate biosynthetic pathway and native parts of yeast metabolism: PPP – pentose phosphate pathway, DAHP – 3-deoxy-D-arabinoheptulosonate 7-phosphate, 3-DHS – 3-dehydroshikimic acid, Trp – tryptophan, Phe – phenylalanine, Tyr – tyrosine. Key enzymes of the shikimate pathway: ARO3 and ARO4 – DAHP synthase isoenzymes and ARO1 – pentafunctional enzyme catalyzing DAHP conversion to chorismate.

Figure 2. Schematic representation of VG pathway reconstruction in two different *S. cerevisiae* background strains. T1, T2 – terminators of *ADH1* and *CYC1*, respectively; P1, P2 – promoters of *PGK1* and *TEF1*, respectively. Red arrows represent essential genetic elements of *S. cerevisiae* and numbered yellow boxes are integration sites characterized by Mikkelsen et al [13].

Figure 3. Growth profiles of the two strains in the batch cultures: a) S-VG and b) C-VG. Grey areas represent ethanol respiration (ER) phase; GF – glucose fermentation. qCO_2 – carbon dioxide production rate, qO_2 – oxygen consumption rate; Glc – glucose, Eth – ethanol and DW – cell dry weight. Error bars represent standard deviation, n=2.

Figure 4. Final yields (45 hours after inoculation) of VG, its intermediates, and TCV produced by S-VG (blue) and C-VG (yellow) during batch cultivation. Error bars represent standard deviation, n=2.

Figure 5. Metabolic profiles of VG and its intermediates in batch cultivation in a) S-VG and b) C-VG strains. c) Accumulation of VG pathway metabolites during the glucose fermentation (GF) and ethanol respiration (ER) growth phases.

Figure 6 Distribution of VG pathway metabolite per biomass throughout entire batch cultivation: a) TCV, b) PAC, c) VAC, d) VG. Blue circles data for S-VG, yellow circles – C-VG. Grey areas represent ethanol respiration phase (narrow in C-VG and wide in S-VG). Error bars represent standard deviation, n≥2. The amounts of PAL and VAN constituted less than 7% percent of TCV at all time points and were not depicted.

Figure 7. Yields of VG, its intermediates, and TCV produced by S-VG (blue) and C-VG (yellow)during continuous cultivation; a) Average of last three samples with more than one retention time in between. Error bars represent standard deviation, n=3.

Supplementary Information

Table S1. Yields ($Y_{S\ Met}$ in mg/g(Glc)) and production rates of VG and its intermediates (r_{Met} in C-mmol/g(DW)·h). Errors represented as a standard deviation, 2≤n≤3.

	Batch [A]		**Chemostat**	
Strain	S-VG	C-VG	S-VG	C-VG
$\mathbf{Y}_{S\ PAC}$	19.0 ± 1.8	15.2 ± 0.2	3.57 ± 0.62	0.26 ± 0.04
$\mathbf{Y}_{S\ PAL}$	1.9 ± 0.2	2.0 ± 0.2	ND	0.09 ± 0.01
$\mathbf{Y}_{S\ VAC}$	1.5 ± 0.2	4.3 ± 1.5	2.56 ± 0.66	3.65 ± 0.10
$\mathbf{Y}_{S\ VAN}$	0.1 ± 0.1	0.2 ± 0.1	0.18 ± 0.14	ND
$\mathbf{Y}_{S\ VG}$	**5.4** ± 0.8	**7.4** ± 0.3	**13.7** ± 1.4	**1.06** ± 0.18
$\mathbf{r}_{PAC}$	1.59 ± 0.35	1.03 ± 0.03	0.64 ± 0.11	0.044 ± 0.008
$\mathbf{r}_{PAL}$	0.18 ± 0.04	0.15 ± 0.01	ND	0.016 ± 0.002
$\mathbf{r}_{VAC}$	0.13 ± 0.03	0.31 ± 0.10	0.46 ± 0.12	0.62 ± 0.03
$\mathbf{r}_{VAN}$	0.01 ± 0.01	0.01 ± 0.01	0.17 ± 0.12	ND
$\mathbf{r}_{VG}$	**0.45** ± 0.12	**0.50** ± 0.02	**2.47** ± 0.25	**0.18** ± 0.03

Note: A – for the batch experiments calculations were made only for exponential growth phase on glucose.

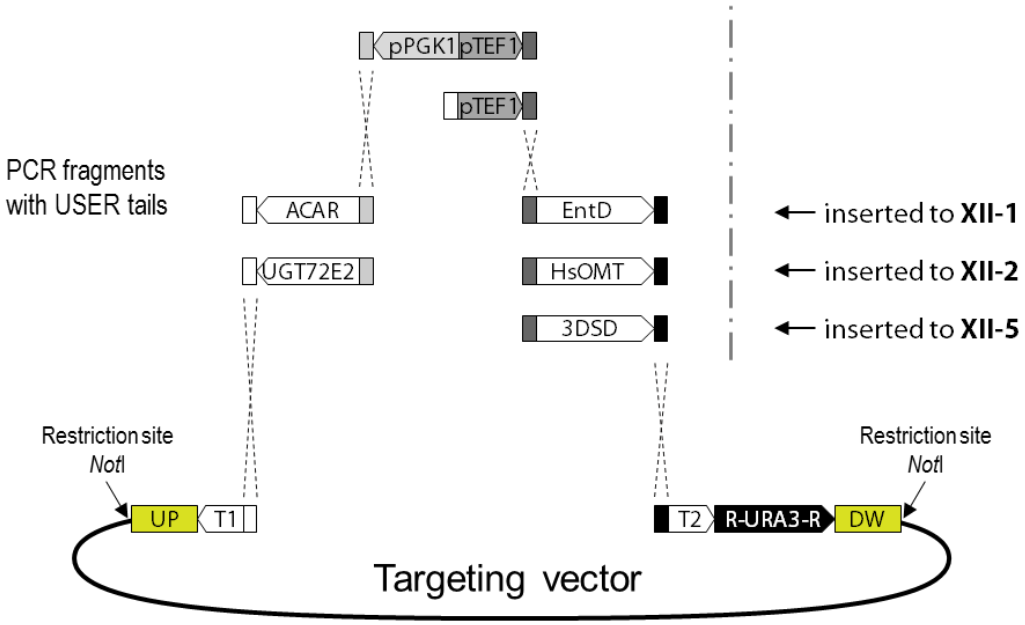


Figure S1. Schematic representation of cloning procedure for transferring vanillin-β-glucoside pathway genes. Targeting vectors pXII1, pXII2 and pXII5 are linearized with AsiSI and Nb.BsmI enzymes. Yellow boxes UP and DW are targeting sequences homologous to chromosomal integration sites. T1 and T2 are terminator sequences Tadh1 and Tcyc1, respectively. Black arrow is counter-selectable auxotrophic marker URA3. White, grey and black boxes represent USER cloning compatible tails.

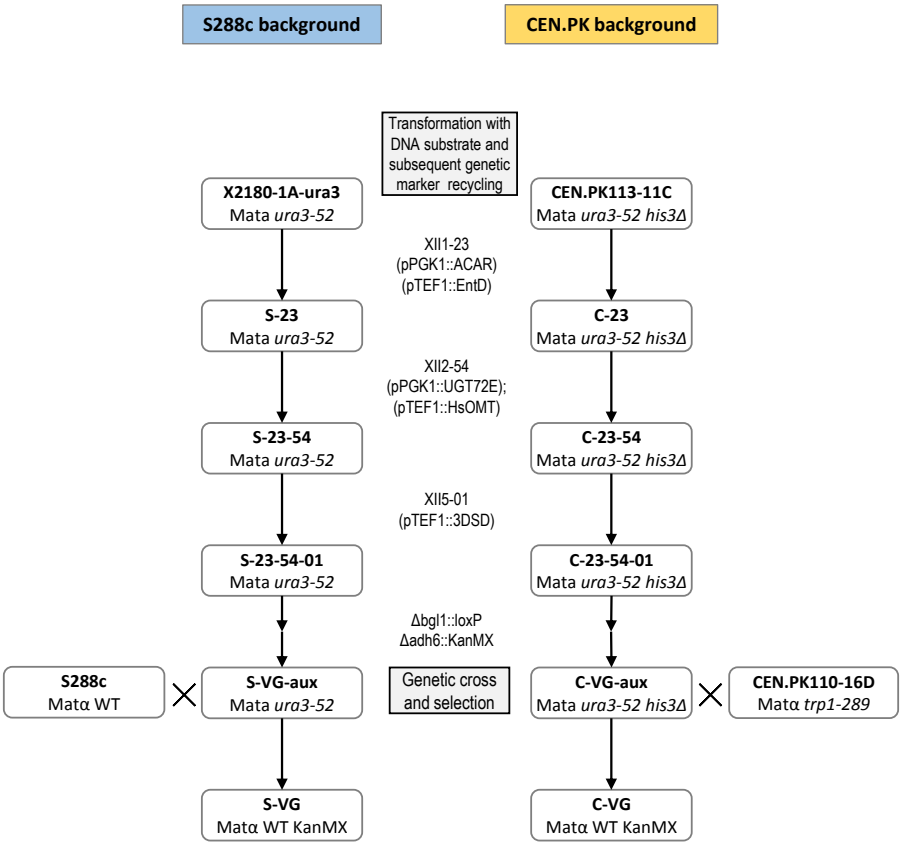


Figure S2 Flowchart representing reconstruction of Vanillin-β-glucoside producing strains. De novo biosynthetic pathway was reconstructed in two strains by sequential transformation with appropriate DNA fragments depicted in the center of the figure. Native genes were deleted with *loxP*-KanMX-*loxP* cassette. Prototrophy was restored by genetic cross.

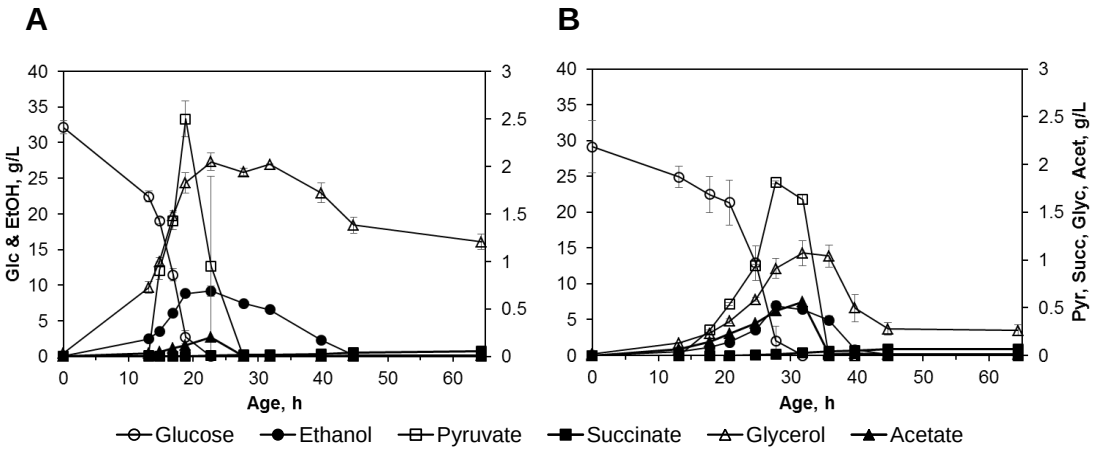


Figure S3 Primary metabolite profiles in strains S-VG (a) and C-VG (b) during batch cultivation. (Error bars represent standard deviation, n=2).

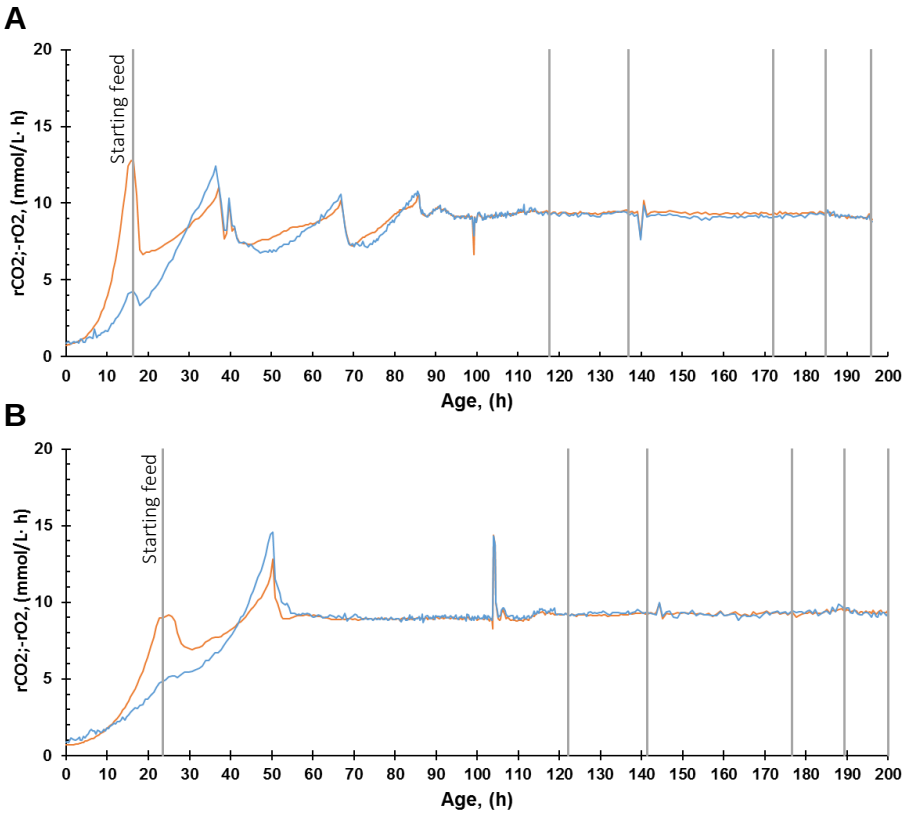


Figure S4 Time profiles of online fermentation data of two strains A) S-VG and B) C-VG in chemostat cultivation mode. Orange line – volumetric rates of CO_2 production; red line – volumetric rates of oxygen consumption. Grey vertical lines represent the sampling points and the start of the chemostat phase. One residence time is equal to 10 hours.

6.5 Paper 5 - LC-MS/MS quantification of intracellular metabolites in *Saccharomyces cerevisiae* using ^{13}C-labeling to minimize matrix interference.

Olivera Magdenoska, Peter Boldsen Knudsen, Daniel Killerup Svenssen, Kristian Fog Nielsen.

Paper submitted to Analytical Biochemistry 2015

LC-MS/MS quantification of intracellular metabolites in *Saccharomyces cerevisiae* using ^{13}C-labeling to minimize matrix interference

Olivera Magdenoska[1,*], Peter Boldsen Knudsen[1], Daniel Killerup Svenssen [1, 2], Kristian Fog Nielsen[1]

[1]Eucariotic Biotechnology, Department of Systems Biology, Technical University of Denmark, Søltofts Plads 221, DK-2800 Kgs. Lyngby, Denmark.

[2]Novo Nordisk Foundation Center for Biosustainability, Kogle Allé 4, 2970 Hørsholm Denmark

[*]Corresponding author. Tel: +45 45252725; Fax: + 45 45884148; E-mail: olima@bio.dtu.dk

Subject category: Mass spectrometry

Short title: Quantification of intracellular metabolites

Abstract

For quantification of intracellular metabolites, LC-MS/MS is currently the method of choice, especially when combined with stable isotopically labeled internal standards (SIL-IS). However due to the presence of the intracellular metabolites in the biological matrix a standard addition based analytical validation is needed. Here we present an alternative solution for minimizing the signal interference from the biological matrix on both the analytes and their SIL-IS's and for compounds with more than 12 C atoms the interference is totally eliminated. This was done by using a matrix obtained from a cultivation of *Saccharomyces cerevisiae* in [$^{13}C_6$]glucose/non-labeled glucose (50/50, w/w) growth medium. The areas of both $^{12}C_6$ and $^{13}C_6$ fractions of ATP in the matrix were measured to be 2 % of the sum of the areas of all ATP isotopes detected in the matrix. The produced biological matrix allowed spiking of both the non-labeled and the SIL-IS and thus more straightforward method validation. The intra and inter day accuracy and precision were ≥80% and ≤20%, respectively. The methodology was used for quantification of nucleotides, coenzymes and redox compounds from *Saccharomyces cerevisiae*. The determined energy charge ratio was 0.9 while the Mal-CoA/Ac-CoA ratio was 0.04.

Keywords: intracellular metabolites; ion-pair UHPLC-MS/MS; blank matrix; ^{13}C labeling; validation; stable isotopically labeled internal standards (SIL-IS).

Introduction

The intracellular metabolome does not only supply the cell with energy but also provides building blocks, and is involved in the cell signaling pathways [1-3]. Determining the intracellular metabolite pool sizes thus helps in defining the metabolic state of the cell and in identifying the bottlenecks in precursor supply as possible targets for metabolic engineering. However, most of intracellular metabolites are ionic and highly polar providing a formidable analytical task seen from a separation perspective. Even more importantly, the enzyme activities need to be quenched instantaneously to provide the true metabolic image of the cell which is often done by changes in temperature or pH [4-7]. Furthermore, the extraction methods are often not selective and result in a very complex biological extracts that contain large amounts of different interfering compounds and salts that can impair the analysis of the metabolites of interest. Therefore both sensitive and very selective analytical methods are required with liquid chromatography tandem mass spectrometry (LC-MS/MS) or capillary electrophoresis MS/MS (CE-MS/MS) as the most commonly applied techniques [8-9]. For the ionic and polar intracellular compounds, electrospray (ESI) ionization is needed which unfortunately suffers significantly from suppression and/or enhancement of the signal [10-11]. Although the exact reasons for the signal modulation are still debated, several theories exists such as: i) formation of more $[M+Na]^+$, $[M+H+Y]^+$ and $[M+X]^+$ rather than $[M+H]^+$ (where Y is a neutral and X a charged compound) in positive mode, while in negative mode $[M+Na-2H]^-$, $[M+X]^-$ and $[M+Y-H]^-$ rather than $[M-H]^-$ could be formed [12], ii) increased surface tension of the droplets during evaporation in the ESI [13] and iii) limited amount of charges available in the ESI droplets [14]. Altogether these factors result in very different signals for the same analyte concentration in different biological matrices as well as from a clean solution, leading to over- or underestimation of the concentration of analytes and thus incorrect interpretation of the metabolomics data. The use of a stable isotopically labeled internal standard (SIL-IS) is the optimal solution to the problem with the LC-MS signal modulation as a result of the matrix interferences as they both can compensate for losses during sample preparation and elutes at the same time as the analyte, thus suffering the same spray modulation [5, 15].

Despite of the approach chosen, the quantification of the metabolite of interest requires that the response of the calibrants correspond to the response of the metabolite of interest in the true analyte-free matrix. However, it is often difficult if not impossible to find a matrix free of the analyte, thus complicating the quantification [16,17]. Furthermore, the analytical validation is becoming increasingly important, but the absence of guidelines for validation of analytical methods for intracellular metabolites together with the absence of a certified analyte-free matrix makes the validation less straightforward [16,17]. This article describes a method for quantification of intracellular metabolites in *Saccharomyces cerevisiae* (*S. cerevisiae*) based on SIL-IS UHPLC-MS/MS. The method has been validated using a novel approach for producing [$^{13}C_6$] labeled/non labeled (50/50, w/w) biological matrix, where the interference is minimized with respect to both the signal from the unlabeled naturally occurring analytes spiked in and to the SIL-IS. The described approach eliminates the need for double labeled (e.g. ^{13}C and ^{15}N) SIL-IS's that are very costly and often unavailable.

Methods and material

Chemicals

The chemicals for Yeast extract-Peptone-Dextrose (YPD) and minimal medium were obtained from Merck Millipore (Darmstadt, Germany) or Sigma-Aldrich Co. (St. Louis, MO/USA). Standards of ATP, ADP, AMP, NAD^+, NADH, NADPH, $NADP^+$, acetyl coenzyme A (Ac-CoA) and malonyl coenzyme (Mal-CoA), [$^{13}C_{10}$]ATP and [$^{15}N_5$]ADP were purchased from Sigma–Aldrich (Steinheim, Germany) while [$^{15}N_5$]AMP was from Silantes (Munich, Germany). [$^{13}C_6$]glucose [$^{13}C_3$]Mal-CoA, and [$^{13}C_2$]Ac-CoA were from Euriso-top (Gif sur Yvette Cedex, Paris, France). Tributylamine (TBA) (puriss plus grade), methanol, CH_3Cl, 1,4-Piperazinediethanesulfonic acid (PIPES), EDTA and acetic acid (LC–MS grade) were obtained from Sigma–Aldrich. Glucose test strips were purchased from Machery-Nagel, Düren, Germany. Water was purified using a Milli-Q system (Millipore, Bedford, MA, USA).

Stock solutions

Stock solutions of the unlabeled compounds with concentration of 1 mg/ml were prepared in H_2O and stored under –80 °C until use. Aliquots of the stock solutions were used to prepare the daily working sub stock solutions used in the quantification and the validation by further dilution in 10 mM TBA and 10 mM acetic acid solution. Two sub stock solutions were used: one containing 325 µg/ml ATP, 130 µg/ml of ADP, AMP, NADH and NADPH and the other one containing 130 µg/ml of Ac-CoA and Mal-CoA and 91 µg/ml of NAD^+ and $NADP^+$ (keeping the redox pairs separated for assessment of the degree of oxidation/reduction during the sample preparation).

Stock solutions (1 and/or 0.2 mg/ml) of SIL-IS were prepared in H_2O and kept under –80 °C until use. Aliquots of these stock solutions were used to prepare a mixture containing 5 µg/ml of [$^{13}C_{10}$]ATP, [$^{15}N_5$]ADP, [$^{15}N_5$]AMP and 10 µg/ml of both [$^{13}C_3$]Mal-CoA and [$^{13}C_2$]Ac-CoA. A volume of 40 µl from this mixture, was used to spike both the standards and the samples.

Preparation of ^{13}C labeled Saccharomyces cerevisiae extracts

Saccharomyces cerevisiae CEN.PK113-7D (MATα, MAL2-8c SUC2) was maintained on YPD medium. Shake flask cultivations were carried out in defined medium [18] with 20 g.L-1 D-glucose (non-labeled for pre-cultures while for isotope enrichment experiments [$^{13}C_6$]glucose/non-labeled glucose (50/50, w/w)). The inoculum for the isotope enrichment experiment was prepared by sub-culturing a single colony from a plate in a shake flask with 100 ml of sterile minimal medium for 14-16 h at 30 °C at 150 rpm. From this exponentially growing pre-culture, cells were harvested and used to inoculate shake flasks containing 100 ml of sterile isotopically enriched medium to a final optical density (OD) of 0.001. The entire culture (OD ~4) was harvested after approx. 22 h, while it was still exponentially growing determined by the presence of glucose measured by glucose test strips.

At OD ~4, 5 ml of the *S. cerevisiae* culture was quenched in 20 ml of MeOH/H_2O (60/40, v/v) precooled to -40 °C in an ethanol/dry ice bath followed by centrifugation at -10 °C for 3 min at 5000 x g. After decanting, the pellets were re-dissolved in ice-cold buffer (pH 7.2) containing 3 mM PIPES and 3 mM EDTA. Two milliliter of the re-dissolved pellet was extracted by adding 7.5 ml of MeOH:CH_3Cl (1:2) mixture. The reason for adding the buffer before the MeOH:CH_3Cl was to avoid the formation of the pellets and ability to split the labeled matrix in equal portions further used in the validation and the quantification. The samples were shaken in a 4 °C room for 60 min followed by centrifugation for 3 min at 4 °C and 5000 x g in order to separate the organic and the H_2O phase. The upper H_2O phase, 3.7 ml was transferred to another 15 ml falcon tube. To the rest of the MeOH:CH_3Cl biomass mixture, 2 ml of the PIPES EDTA buffer pH 7.2 and 2 ml MeOH was added followed by mixing. The samples were once more centrifuged using the above conditions followed by collecting 4 ml of the upper phase that was combined with the 3.7 ml taken from the first centrifugation. After evaporation using lyophilisation, the samples were re-dissolved in 200 µl of eluent A and filtered through a hydrophilic PTFE filter (Advantec, Tokyo, Japan, diameter 13 mm; pore size 0.2 µm).

Preparation of unlabeled Saccharomyces cerevisiae extracts

Chemostat cultivations were carried out in Sartorious and 1 l bioreactors (Satorious, Stedim Biotech, Goettingen, Germany) with a working volume of 0.6 l, equipped with baffles and 2 Rushton six-blade disc turbines. Culture conditions were pH 5, stir rate 800 rpm and air flow 1 volume of air per volume of liquid per minute (vvm). The pH electrode (Mettler Toledo, OH/USA) was calibrated according to manufacturer's standard procedures. The bioreactor was sparged with sterile atmospheric air and off-gas concentrations of oxygen and carbon dioxide were measured with a Prima Pro Process Mass Spectrometer (Thermo-Fischer Scientific, Waltham, MA/USA), calibrated monthly with gas mixtures containing 5 % (v/v) CO_2, 0.04 % (v/v) ethanol and methanol, 1 % (v/v) argon, 5 % (v/v) and 15 % (v/v) oxygen all with nitrogen as carrier gas (Linde Gas, AGA, Enköping, Sweden). Temperature was maintained at 30 °C throughout the cultivation and pH controlled by automatic addition of 2 N NaOH and H_2SO_4.

Continuous cultivations were initiated and cultivated as batch until late exponential phase where supply of the feed medium was initiated ($D = 0.1\ h^{-1}$). The feed medium was identical to the batch medium with the exception of the glucose concentration, which was 10 g/l glucose. Feed supply was controlled by a gravimetrically controlled peristaltic pump and reactor volume was kept constant applying a level probe ensuring continuous removal of excess volume. Automatic addition of feed was initiated at late exponential growth at a rate of 60 ml per hour ensuring a dilution rate $D = 0.1\ h^{-1}$. The culture was left for at least 5 residence times before samples were taken ensuring steady state conditions.

The quenching and the extraction were done as previously explained for the labeled biomass. The only difference was that during the extraction first 7.5 ml of MeOH:CH_3Cl (1:2) were added to the biomass followed by the addition of 2 ml of the PIPES EDTA buffer.

Ion-pair UHPLC

All experiments were carried out using Agilent 1290 binary UHPLC system. The chromatographic conditions have previously been described in Magdenoska et al. [5]. In brief ion pair chromatography was used for separation using TBA as ion pair reagent. The gradient used was: 19.5 min. (0-12 min.

from 0 to 50 % B, 12-12.5 min. from 50 to 100 % B, 12.5 -14 min. 100 % B, 14-14.5 min. from 100 to 0 % B, 14.5-19.5 0 % B). The injection volume used was 3 µl.

Full scan and MS/MS measurements

Agilent 6460 triple quadrupole system (Torrance, CA, USA) equipped with an Agilent Jet Stream ESI source was used for the validation and quantification. The optimized ion source MS parameters used in this study has previously been described in Magdenoska et al. [5]. The MS was operated in negative ion mode with nitrogen as a collision gas. For full scan experiments the MS was operated in MS2 Scan mode and spectra were acquired in the range between m/z 50 and 1000. Multiple reactions monitoring (MRM) was used for the validation and quantification experiments. Both a quantifier and a qualifier ion were monitored for each metabolite. To verify the metabolite presence in the samples, a qualifier/quantifier ion ratio within ± 20% was required (Table 1).

For the high resolution MS measurements, the ion-pair effluent from the Agilent 1290 binary UHPLC system was passed through an Agilent 6550 quadrupole time-of-flight (QTOF) instrument operated in negative and the 2-GHz extended dynamic mode at a resolution of 25,000 full width at half-maximum (FWHM).

Linearity, quantification and limit of detection (LOD)

To check the linearity, 2 ml of the previously prepared isotope enriched biomass, re-dissolved in the PIPES/EDTA buffer, was pipetted into 20 50 ml spin tubes, followed by addition of 7.5 ml MeOH:CH_3Cl (1:2, v/v) extraction mixture. Fourteen of these tubes were spiked with 40 µl of IS mix and different volumes of the previously prepared stock solution in order to get concentration levels ranging from 2.5 to 250 µg/ml for ATP, 1 to 100 µg/ml for AMP, ADP, NADPH, NADH, NAD^+, Ac-CoA and Mal-CoA and 0.7 to 70 µg/ml for $NADP^+$, NAD^+. The remaining six tubes were kept as matrix blanks. Three of the matrix blanks were spiked with 40 µl of IS mixture and the other three were spiked only with H_2O to simulate the spiking of the standards. The samples were processed through the extraction procedure described above.

For ATP, ADP, AMP, Ac-CoA and Mal-CoA for which SIL–IS was available, the calibration curves were prepared by plotting the ratio between the peak areas of the non-labeled standard and SIL-IS versus the concentration of the particular compound. For the redox compounds for which SIL-IS was not available the calibration curves were prepared by plotting the peak area of the standard versus the concentration. Six matrix blanks were analyzed during the quantification and the average peak areas of each metabolite in the matrix blanks was subtracted from the peak area of that metabolite found in the spiked samples to correct for naturally-occurring amount of the particular metabolite [8]. Weighted linear regression (weighted to 1/x) was used to establish the linearity for each of the compounds. The generated calibration curves were also used for quantification of the non-labeled biomass. The calculations were done in Excel.

The limit of detection was estimated to be the lowest concentration at which the signal was at least 5-fold of the corresponding background.

Matrix effects

To determine the matrix effect over the gradient a T-piece infusion test was conducted. While infusing 10 µg/ml ATP into the MS at a constant flow using a syringe pump, two samples were injected: i) eluent A and ii) blank matrix processed through the sample preparation procedure.

Quantitative determination of the matrix effects was done using the Matuszewski et al. approach. Two ml of the previously prepared biomass was pipetted into 14 50 mL spin tubes, followed by addition of 7.5 ml of MeOH:CH_3Cl (1:2, v/v). To each of the tubes, 200 µl of H_2O was added in order to simulate the spiking of the standards. The samples were processed through the extraction. After the evaporation by lyophilisation, the pellet was re-dissolved in the 40 µl of IS and a volume of the previously prepared stock solutions in order to get the above given concentration ranges for the different compounds followed by filtration.

Overall process recovery

The process recovery was inspected at three different concentration levels in duplicates for all the metabolites by using the method described in Canelas et.al. [19]. 2 ml of the matrix was pipetted in 18

50 ml spin tubes resulting in 2 replicates per level. Half of them were spiked before the extraction with 40 µl of IS mix and a volume of the stock solution containing all the metabolites so that the final concentration after evaporation and re-dissolving is 2.5, 7.5 and 175 µg/ml for ATP, 3, 10 and 70 µg/ml for ADP, AMP, $NADP^+$, NADPH, Mal-CoA and Ac-CoA and 2.1, 7 and 49 µg/ml for NAD^+ and $NADP^+$. The other half was spiked only with the standard mix. All the spiked samples were processed through the extraction procedure. After the extraction and the evaporation the samples that contained IS were re-dissolved in 200 µl of eluent. The rest of the samples were re-dissolved in 40 µl of IS and a volume of NAD^+ and $NADP^+$ or NADH and NADPH solution, so that the final concentrations for a particular metabolite after refilling to 200 µl with eluent A corresponds to the ones given above. The difference between the samples spiked before and after extraction with IS and NAD^+/$NADP^+$ or NADPH/NADH will give information on the losses during the extraction procedure.

Accuracy/recovery and precision

The intra and inter day variability of the method were inspected by using the same concentration levels as for the overall process recovery in triplicates on three different days. 2 ml of the previously prepared biomass was pipetted into 18 50 ml spin tubes, followed by addition of 7.5 ml of MeOH:CH_3Cl (1:2, v/v) extraction mixture. The tubes were spiked with 40 µl of IS mix and different volumes of the previously prepared stock solution in order to get concentration levels corresponding to those for the inspection of the process recovery. The samples were extracted using the extraction procedure described above.

Results and discussion

Instrumental method

Ion-pair UHPLC was coupled to MS/MS instrument for the analysis of ATP, ADP and AMP as well as NAD^+, $NADP^+$, NADH, NADPH, Ac-CoA and Mal-CoA. The precursor and product ions, as well as RT used in the MS/MS method are given in Table 1.

[Table 1]

Using a 19.5 min, gradient successful baseline separation of all metabolites was obtained, except for the NADPH and the coenzymes which were coeluting. However, due to the difference in their precursor ions, separation via MS was possible, except for Ac-CoA and its SIL-IS. The reason for this was the 2 Da difference in their masses, which resulted in signal contribution from the [A+2] isotopomer from the non-labeled Ac-CoA to the SIL-IS. As previously also reported by Tan et al. [20] the cross signal contribution resulting from isotopic interference was not expected to cause any systematic errors since the same internal standard was used to prepare the calibration curves, the validation samples and the real samples. The % bias (data not shown) of the validation samples prepared using 10 µg/ml SIL-IS of the Ac-CoA was within the acceptance criteria of ±20%. Figure 1 shows the chromatograms of the labeled and non-labeled compounds.

[Figure 1]

Matrix used for validation and quantification

The main purpose of this study was to explore alternative ways of producing a biological matrix with negligible influence to the MS/MS signals of the metabolites of interest and their SIL-IS, while still containing all the endogenous compounds and enzymes. For this purpose *S. cerevisiae* was cultivated in medium containing [$^{13}C_6$]glucose/non-labeled glucose (50/50, w/w). This ratio of labeled and non-labeled glucose for cultivating the cells resulted as expected in pools of metabolites with labeling in different carbon positions. Figure 2 shows an isotopic pattern of ATP measured in the extract from the *S. cerevisiae* cultivated using this glucose mixture. It should be noted that unless otherwise stated, further on in the text the term ‘’blank matrix’’ refers to the extract obtained from the *S. cerevisiae* cultivated on this mixture.

[Figure 2]

From the measured and calculated isotopic pattern of ATP shown in Figure 2, it can be seen that the intensity of the peaks that correspond to m/z 505.9 ATP and m/z 515.9 [$^{13}C_{10}$]ATP, detected in the matrix, were very low (approximately 10 % of the highest isotope detected). The area of the [$^{12}C_6$]ATP and [$^{13}C_6$]ATP detected in the blank matrix were 13 % and 10 % relative to the area of the most abundant isotope detected (abundance of 100%), respectively. Consequently, there was a small

298 interference of the naturally occurring ATP and [$^{13}C_{10}$]ATP in matrix to the spiked amount of ATP
299 and [$^{13}C_{10}$]ATP, respectively. The highest interference from the matrix was to the MS signals of
300 [$^{15}N_5$]ADP and [$^{15}N_5$]AMP while the lowest was to the coenzymes and the redox compounds. For
301 comparison Figure 3 shows the normalized chromatographic peaks of [$^{15}N_5$]ADP, [$^{15}N_5$]AMP, ATP
302 and Mal-CoA in the blank and spiked matrix, respectively, both processed as specified in the sample
303 preparation.

304 **[Figure 3]**

305 As expected, increase in the molecular mass of only 5 Da was observed for the ^{15}N labeled ADP and
306 AMP. This increase was also detected in the biological matrix obtained by growing *S. cerevisiae* in
307 media containing [$^{13}C_6$]glucose/non-labeled glucose (50/50, w/w) (Figure 3), thus choosing ^{13}C
308 instead of ^{15}N labeled ADP and AMP in this context would have resulted in less interference.
309 However, [$^{13}C_{10}$]ADP and [$^{13}C_{10}$]AMP were commercially not available when the study was
310 conducted. Furthermore, due to the inconsistency in the labeling (all ^{13}C, ^{15}N or $^{13}C^{15}N$) of the
311 commercially available SIL-IS for the different compounds, it is difficult to produce a labeled matrix
312 that will exhibit no interference to the SIL-IS signal. The amount in percentage of the non-labeled
313 compounds that were replaced by their ^{13}C fractions detected in the matrix is given in Table 2.

314 **[Table 2]**

315 The amount of some of the non-labeled metabolites present in the blank matrix was calculated to be
316 approximately 50 times lower when compared to the amount expected to be measured if *S. cerevisiae*
317 was grown on non-labeled glucose. Finally, the fewer the number of the carbons in the molecule, the
318 higher the interference. Theoretically, the ^{12}C fraction of a compound that has 15 carbons will have 30
319 x lower abundance than the ^{12}C fraction of a compound with 10 carbons. This shows the suitability of
320 the presented approach for generating blank matrices with low background amounts of the compounds
321 of interest.

322

323 *Validation*

324 The validation of the methods was done by determining: i) linearity; ii) the overall process recovery;
325 iii) matrix effects; iv) intra day and inter day accuracy/recovery and precision.

Linearity. The linearity of the method was inspected by injecting 3 µl of samples prepared by spiking the matrix with standards, with concentrations over three orders of magnitudes, and subsequently processed through the sample preparation as explained in the Methods and material part. Different span of working concentrations were chosen due to the different amounts of the various compounds present in the cell. In general, all investigated metabolites showed good linearity over the inspected concentration range with correlations coefficients (R^2) of ≥0.990 (Table 3).

[Table 3]

Due to the cross signal contribution for Ac-CoA, resulting from isotopic interference, several concentrations (1, 3, 5 and 10 µg/ml) of internal standard were tested, with 10 µg/ml giving the best linearity (R^2 of 0.998). The need for higher SIL-IS concentration for improved linearity when faced with cross signal contribution from isotopic interference has previously been reported by Tan et al. [20].

Overall process recovery. For ATP, ADP, AMP, Mal-CoA and Ac-CoA, for which SIL-IS were available, the overall process recovery was determined according to Canelas et al. [19] at three different concentrations. Two types of samples were prepared for each of these compounds in duplicates. All samples were spiked with standard mixture before the extraction, whereas only half of them were spiked with SIL-IS mixture before the extraction and the other half before the analysis. The recovery was calculated using the following formula [1]:

$$Recovery\ when\ SIL-IS\ was\ avaliable = \frac{c_i(BA)}{c_i(BE)} \quad [1]$$

where c_i *(BA)* is the determined concentration of the analyte *i* in the sample prepared by spiking the blank matrix with the standard analyte *i* with certain concentration, processing the sample through the extraction and spiking it with the corresponding SIL-IS before the analysis, while c_i *(BE)* is the determined concentration of the analysis *i* in the sample prepared by spiking the blank matrix with the standard analyte *i* and the corresponding SIL-IS and processing it through the extraction procedure.

For the compounds, for which no internal standard was available, the recovery was determined by spiking half of the samples with non-labeled standards before the extraction and the other half before the analysis. The recovery was determined using the following formula [2]

[2]

353 $$Recovery\ when\ SIL-IS\ is\ \boldsymbol{not}\ avaliable = \frac{c_i(BE)}{c_i(BA)}$$

354 The overall process recovery of all of the metabolites investigated during this study is shown in
355 Figure 4, where the process recovery is plotted against the compounds investigated.

356 **[Figure 4]**

357 As shown in Figure 4, the adenosine mono-, di- and tri-phosphates, the coenzymes and the oxidized
358 forms of the redox pairs showed a recovery of 0.7-0.8, indicating that high percentage of the
359 metabolites investigated, has been recovered during the sample preparation procedure. Slightly lower
360 recovery of 0.6 was obtained for the NADH.

361 Matrix effects. The presence of ion suppression or enhancement was evaluated both qualitatively and
362 quantitatively by using two approaches: i) the T-piece infusion test [21] and ii) the approach by
363 Matuzewski et al. [10, 15], respectively. In order to evaluate the level of suppression or enhancement
364 of the MS signal over the LC gradient, the T-piece infusion test was performed. While having a
365 constant infusion of 10 µg/ml ATP into the MS, two samples were injected: matrix that has been
366 processed through the extraction procedure and eluent A. Since ATP is constantly infused into the MS
367 with a stable ion response over time, any compound that elutes from the column and goes into the MS
368 will cause suppression or enhancement of the signal from the infused ATP. Figure 5 shows the
369 chromatograms of post column infusion of 10 µg/ml ATP into the MS after injecting eluent A or
370 matrix blank superimposed with the chromatograms of the compounds of interest to show their
371 retention times.

372 **[Figure 5]**

373 In general the highest suppression was observed at the beginning and in the middle of the gradient.
374 The suppression at the beginning was a result of not well retained compounds that elute in the time
375 segment at or near the void volume of the column. The observed suppression in the middle of the
376 gradient was due to the matrix compounds that have been retained on column and elute later in the
377 analysis. The drop of the signal in the region between 7 and 8 min. and between 11 and 12 min. was
378 due to the elution of the two buffering reagents used for the sample preparation procedure: PIPES and
379 EDTA, respectively. PIPES formed dimer, trimer, tetramer and pentamer that were clearly seen in the

full MS scan with a mass difference between the aggregates of m/z 302.0. In the region where AMP eluted high intensity peak that had an m/z 96.9 and corresponds to a phosphate was detected, causing drop of the signal intensity as well. The strongest ion enhancement was mostly present at the end of the gradient in the region where none of the metabolites of interest eluted. Since the degree of suppression or enhancement is compound and concentration dependent, Matuzewski et al. [10] approach was adopted to investigate how does the matrix affect the compounds LC-MS signals at both high and low concentrations. The calculations from the quantitative evaluation of the matrix effects have been summarized in Table 4.

[Table 4]

All compounds exhibited suppression of the signal (36-91 %) and lower values for the slope of the matrix matched calibration curves when compared to the slope of the calibration curves obtained from pure standards. The strongest suppression was observed for AMP (91 %) and ADP (89 %), due to the elution of phosphate and EDTA, respectively.

For inspecting whether the matrix effects and the losses during the sample preparation were corrected by using the internal standard, two calibration curves prepared over three days were compared: i) in the neat solution, ii) and the standard curves obtained by spiking the matrix before the extraction. For the compounds for which SIL-IS was available, the linear regression equations of the calibration curves obtained using the standards in the neat solution and the ones in the pure standards were comparable (data not shown). This indicates that the calibration curves prepared in the neat solution may also be used for quantification with this type of matrix. However, we suggest using the matrix spiked and processed standard curves, since SIL-IS is not always available, and thus reducing the number of samples per run.

Accuracy/recovery and precision. The intra day precision and accuracy/recovery were determined by using matrix spiked QC samples prepared in triplicates at three different concentration levels processed in accordance with the sample preparation procedure. The inter day precision and accuracy/recovery were evaluated over three different days. The precision was given as the relative standard deviation of the determined concentrations, while the accuracy was calculated by dividing the mean measured concentration by the nominal concentration times 100 [22]. The results from the

validation (Table 3) showed acceptable precision and accuracy based on the commonly accepted criteria for a quantitative method (RSD less than 20 % and accuracy within 80-120 %).

Evaluation of oxidation of NADH and NADPH to NAD^+ and $NADP^+$ respectively

NADH and NADPH can be oxidized to NAD^+ and $NADP^+$, respectively, during the sample preparation. Therefore obtaining a snapshot of the levels of the NAD^+/NADH and $NADP^+$/NADPH can be very challenging. However during this study even oxidation of the NADPH and NADH standards prior to the sample preparation was encountered. Due to the fact that the oxidized and reduced forms of the redox pairs in the spiked samples were kept separated during the sample preparation, it was possible to check the degree of oxidation of NADH to NAD^+ and NADPH to $NADP^+$. The data in Table 5 show how much oxidized species (NAD^+, $NADP^+$) are formed by oxidation of NADH and NADPH. The results given in Table 5 are from triplicates, analyzed on three different days.

[Table 5]

Namely, a peak of NAD^+ was detected in the NADH single standard solution freshly prepared in either H_2O or eluent A. Interestingly enough, in the chromatogram of the single NADH standard, two peaks were detected that had the same mass as NAD^+, where one of them had the same retention time as the NAD^+ standard and the other one was eluting approximately 0.3 min. earlier. By inspecting the fragmentation pattern of these two peaks by interfacing the ion pair LC to a quadrupole time of flight MS (data not shown), it was found that these two peaks have identical fragments, but the ratios between the fragments were different. This leaded to a conclusion that the two peaks might correspond to α and β isomers of NAD^+. In the case of NADPH, only one peak for $NADP^+$ was observed. As previously reported in the literature [23] ammonium acetate pH 8 has been shown to improve the stability of NADPH and reduce the oxidation. In order to investigate this, the standards of NADH and NADPH were prepared in 5 mM ammonium acetate pH 8, which indeed reduced the oxidation of these two compounds. However the addition of TBA in the injection vial was proven to increase the sensitivity of the nucleotides [5]. Therefore 50 μg/ml solutions of NADPH and NADP were prepared by diluting the ammonium acetate stock solutions (pH 8) of these two compounds into

eluent A which contained TBA. It was noted that the addition of TBA did increase the sensitivity especially for NADPH. However further investigations are necessary in order to check the long term stability of the samples prepared in the ammonium acetate as well as the dilutions of these stocks into TBA. Further optimizations of the sample preparation procedure might be necessary as well in order to prevent the oxidation of NADPH and NADH, since it has been shown that the incubation time during CH_3Cl extraction can also have an effect on the extent of oxidation [23].

Application of the method

The validated quantitative method was used to determine the intracellular amounts of adenosine mono-, di- and tri-phosphate, the coenzymes and the redox pairs in *S. cerevisiae*. The quantitative data are given in Table 6 in nmol/mg dry weight. From the measured concentrations of ATP, ADP and AMP the energy charge ratio was determined to be 0.9, indicating a successful quenching and therefore high quality metabolomics data.

[Table 6]

The ratio between Mal-CoA and Ac-CoA was found to be 0.04. Similar concentration of Ac-CoA have also been measured before [24], however according to the authors knowledge values for the Mal-CoA/Ac-CoA ratio in *S. cerevisiae* have not been previously reported.

The determined NAD^+/NADH and $NADP^+$/NADPH ratios were higher than some of the ones reported in the literature [25-30]. The reason for the high ratios was probably the oxidation of NADH and NADPH, consequently an increased concentration of NAD^+ and NADH [30]. However Hou et al. [30] reported data for the NAD^+/NADH and $NADP^+$/NADPH ratios that are close to the ones measured in the present study. It should be noted that the same quenching and extraction procedure were used in the present and the study reported by Hou et al. [31]. Nevertheless, modification of the quenching and extraction procedures might help in prevention of the oxidation of the reduced species. Re-dissolving the samples in ammonium acetate pH 8 instead of eluent A might also be taken into account for improvement of the accuracy of the method. In addition, SIL-IS of the redox compounds, which were not available during this study, will be beneficial for further investigation of the oxidation.

Conclusion

An analytical method for quantitative measurement of intracellular metabolites in *S. cerevisiae* has been established. Using a novel approach, by cultivating *S. cerevisiae* on a media containing [$^{13}C_6$]glucose/non-labeled glucose (50/50, w/w), the interference from the naturally occurring intracellular metabolites in the spiking matrix were removed or reduced to a minimum. The reported approach allowed use of both non-labeled standards and their corresponding SIL-IS, therefore simplifying the method validation and quantification. The results show that the method was suitable for quantifying nucleotides, coenzymes and redox compounds in *S. cerevisiae.* Furthermore, the method can be applied to any analytical samples that exhibit high background amounts of the analytes of interest and when blank matrix free of the analyte is not available.

Acknowledgments

This work was supported by the Danish Research Agency for Technology and Production, grant # 09-064967. The authors acknowledge the Technical University of Denmark Fermentation Platform at The Department of Systems Biology for providing the Agilent 6460 Triple Quadrupole LC/MS instrument utilized in this study. We thank Agilent technologies for the Thought Leader Donation of the 6550 QTOF instrument.

492

493

494

495

496 **References**

497 [1] P. M. R. Guimarães, J. Londesborough, The adenylate energy charge and specific fermentation
498 rate of brewer's yeasts fermenting high- and very high-gravity worts, *Yeast* 25 (2008) 47-58.

499 [2] F. C. Schroeder, Modular Assembly of Primary Metabolic Building Blocks: A Chemical
500 Language in *C. elegans*, *Chem. Biol.* 22 (2014) 7-16.

501 [3] P.W. D'Alvise, O. Magdenoska, J. Melchiorsen, K. F. Nielsen, L. Gram, Biofilm formation and
502 antibiotic production in *Ruegeria Mobilis* are influenced by intracellular concentrations of
503 cyclic dimeric guanosinmonophosphate, Environmental Microbiology 15 (2013) 1252-1266.

504 [4] H. Meyer, H. Weidmann, M. Lalk, Methodological approaches to help unravel the intracellular
505 metabolome of *Bacillus subtilis*, Microbial cell factories 12 (2013) 1-13.

506 [5] O. Magdenoska, J. Martinussen, J. Thykaer, K. Fog Nielsen, Dispersive solid phase extraction
507 combined with ion-pair ultra high-performance liquid chromatography tandem mass
508 spectrometryfor quantification of nucleotides in *Lactococcus lactis*, Anal. Biochem. 4440
509 (2013) 166-177.

510 [6] S. Meinert, S. Rapp, K. Schmitz, S. Noack, G. Kornfeld, T. Hardiman, Quantitative quenching
511 evaluation and direct intracellular metabolite analysis in *Penicillium chrysogenum*, Anal.
512 Biochem. 1104 (2014) 211-221.

513 [7] L. P. de Jonge, R. D. Douma, J. J. Heijnen, W. M. van Gulik, Optimization of cold methanol
514 quenching for quantitative metabolomics of *Penicillium chrysogenum*, Metabolomics 8 (2012)
515 727-735.

516 [8] J. M. Knee, T. Z. Rzezniczak, A. Barsch, K. Z. Guo, T. J.S. Merritt, A novel ion pairing
517 LC/MS metabolomics protocol for study of a variety of biologically relevant polar metabolites,
518 J. Chromatogr. B. 936 (2013) 63-73.

519 [9] Y. Ohashi, A. Hirayam, T. Ishikawa, S. Nakamura, K. Shimizu, Y. Ueno, M. Tomitaa and T.
520 Soga, Depiction of metabolome changes in histidine-starved Escherichia coli by CE-TOFMS,
521 Mol. BioSyst., 4 (2008) 135-147.
522 [10] B. K. Matuszewski, M. L. Constanzer, and C. M. Chavez-Eng, Strategies for assessment of
523 matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS, Anal. Chem. 75
524 (2003) 3019-3030.
525 [11] D. Remane, D. K. Wissenbach, M. R. Meyer and H. H. Maurer, Systematic investigation of ion
526 suppression and enhancement effects of fourteen stable-isotope-labeled internal standards by
527 their native analogues using atmospheric-pressure chemical ionization and electrospray
528 ionization and the relevance for multi-analyte liquid chromatographic/mass spectrometric
529 procedures, Rapid Commun. Mass Spectrom. 24 (2010) 859-867.
530 [12] A. Klitgaard, A. Iversen, M. R. Andersen, T. O. Larsen, J. Christian Frisvad, K. Fog Nielsen,
531 Aggressive dereplication using UHPLC–DAD–QTOF: screening extracts for up to 3000
532 fungal secondary metabolites, Anal. Bioanal. Chem. 406 (2014)1933-1943.
533 [13] C. R. Mallet, Z. Lu and J. R. Mazzeo, A study of ion suppression effects in electrospray
534 ionization from mobile phase additives and solid-phase extracts, Rapid Commun. Mass
535 Spectrom. 18 (2004) 49-58.
536 [14] C.H.P. Bruins, C.M. Jeronimus-Stratingh, K. Ensing, W.D. van Dongen, G.J. de Jong, On-line
537 coupling of solid-phase extraction with mass spectrometry for the analysis of biological
538 samples I. Determination of clenbuterol in urine, J. Chomatogr. A. 863 (1999) 115-122.
539 [15] B.K. Matuszewski, Standard line slopes as a measure of a relative matrix effect in quantitative
540 HPLC–MS bioanalysis, J. Chomatogr. B. 830 (2006) 293-300.
541 [16] N. C. van de Merbel, Quantitative determination of endogenous compounds in biological
542 samples using chromatographic techniques, Trends Anal. Chem. 27 (2008) 924-933.
543 [17] S. Ongaya, G. Hendriksb, J. Hermansa, M. van den Bergec, N. H.T. ten Hackenc, N. C. van de
544 Merbela, R. Bischoffa, Quantification of free and total desmosine and isodesmosine in human
545 urine by liquid chromatography tandem mass spectrometry: A comparison of the surrogate

546 analyte and the surrogate matrix approach for quantification, J. Chomatogr. A. 1326 (2014) 13-
547 19.
548 [18] C. verduyn, E. Postma, W. A. Scheffers and J. P. van Dijken, Effect of benzoic acid on
549 metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and
550 alcoholic fermentation, Yeast 8 (1992) 501-517.
551 [19] A. B. Canelas, A. ten Pierick, C. Ras, R. M. Seifar, J. C. van Dam, W. M. van Gulik, and J. J.
552 Heijnen, Quantitative evaluation of intracellular metabolite extraction techniques for yeast
553 metabolomics, Anal. Chem. 81 (2009) 7379-7389.
554 [20] A. Tan, I. A. Lévesque, I. M. Lévesque, F. Viel, N. Boudreau, A. Lévesque, Analyte and
555 internal standard cross signal contributions and their impact on quantitation in LC-MS based
556 bioanalysis, J. Chomatogr. B. 879 (2011) 1954-1960.
557 [21] T.M. Annesley, Ion suppression in mass spectrometry, Clin. Chem. 49 (2003) 1041–1044.
558 [22] V. Bezy, P. Chaimbault, P. Morin, S. E. Unger, M. C. Bernard, L. A. Agrofoglio, Analysis and
559 validation of the phosphorylated metabolites of two anti-human immunodeficiency virus
560 nucleotides (stavudine and didanosine) by pressure-assisted CE ESI-MS/MS in cell extracts:
561 Sensitivity enhancement by the use of perfluorinated acids and alcohols as coaxial sheath-
562 liquidmake-up constituents, Eleectrophoresis 27 (2006) 2464-2476.
563 [23] K. Ortmayr, J. Nocon, B. Gasser, D. Mattanovich, S. Hann, G. Koellensperger, Sample
564 preparation workflow for the liquid chromatography tandem mass spectrometry based analysis
565 of nicotinamide adenine dinucleotide phosphate cofactors in yeast, J. Sep. Sci. 37 (2014) 2185-
566 2191.
567 [24] R. M. Seifar, C. Ras, A. T. Deshmukh, K. M. Bekers, C. A. Suarez-Mendez, A. L.B. da Cruz,
568 W. M. van Gulik, J. J. Heijnen, Quantitative analysis of intracellular coenzymes in
569 *Saccharomyces cerevisiae* using ion pair reversed phase ultra high performance liquid
570 chromatography tandem mass spectrometry, Anal. Biochem. J. Chromatogr. A, 1311 (2013)
571 115-120.

[25] Lin, S. J., Ford, E., Haigis, M., Liszt, G., Guarente, L., Calorie restriction extends yeast life span by lowering the level of NADH, Genes Dev. 18 (2004) 12-16.

[26] Anderson, R. M., Bitterman, K. J., Wood, J. G., Medvedik, O., Cohen, H., Lin, S. S., Manchester, J. K., Gordon, J. I., Sinclair, D. A., Manipulation of a nuclear NAD+ salvage pathway delays aging without altering steady-state NAD+ levels, J. Biol. Chem. 277 (2002) 18881-18890.

[27] Lin, S. S., Manchester, J. K., Gordon, J. I., Enhanced gluconeogenesis and increased energy storage as hallmarks of aging in Saccharomyces cerevisiae. J. Biol. Chem. 276 (2001) 36000 - 36007.

[28] Ashrafi, K., Lin, S. S., Manchester, J. K., Gordon, J. I., Sip2p and its partner snf1p kinase affect aging in S. cerevisiae. Genes Dev. 14 (2000) 1872-1885.

[29] Smith, J. S., Brachmann, C. B., Celic, I., Kenna, M. A., Muhammad, S., Starai, V. J., Avalos, J. L., Escalante-Semerena, J. C., Grubmeyer, C., Wolberger, C., Boeke, J. D., A phylogenetically conserved NAD+-dependent protein deacetylase activity in the Sir2 protein family, Proc. Natl. Acad. Sci. USA 97 (2000) 6658-6663.

[30] J. L. Sporty, M. M.Kabir, K. W. Turteltaub, T. Ognibene, S.-J. Lin, G. Bench, Single sample extraction protocol for the quantification of NAD and NADH redox states in Saccharomyces cerevisiae, J. Sep. Sci. 31 (2008) 3202-3211.

[31] J. Hou, N. F. Lages, M. Oldiges, G. N. Vemuri, Metabolic impact of redox cofactor perturbations in *Saccharomyces cerevisiae*, Metab. Eng. 11 (2009) 253-261.

Tables

Table 1. Optimized precursor ions, quantifier and qualifier product ions, collision energy, fragmentor and cell accelerator voltage used in the MRM measurements.

Compound	RT (min)	Precursor ion (m/z)	Quantifier				Qualifier (m/z)	Qualifier/Quantifier ion ratio
			(m/z)	CE (V)	Frag (V)	CAV (V)		
ATP	13.4	505.9	273.0	30	95	3	407.9	19.2
ADP	12.0	425.9	134	25	115	3	158.9	83.3
AMP	9.1	346	79	20	130	3	211.1	4.7
Ac-CoA	13.6	808	407.9	40	135	3	460.9	68.7
Mal-CoA	13.6	852	807.9	25	135	4	407.9	49.1
NAD^+	5.4	662	540	15	120	4	407.9	5.6
NADH	10.6	664	397	30	110	4	346	62.7
$NADP^+$	12.5	742	620	15	110	3	408	23
NADPH	13.5	744	408	35	100	4	397	53
$[^{13}C_{10}]$ATP	13.4	515.9	278	30	125	3	417.9	19.4
$[^{15}N_5]$ADP	12.0	431	139.3	20	120	2	158.9	88.6
$[^{15}N_5]$AMP	13.6	350.8	79	35	90	3	211.1	4.7
$[^{13}C_2]$Ac-CoA	13.6	810	407.9	40	135	3	462.8	44
$[^{13}C_3]$Mal CoA	13.6	855	810.2	30	130	4	463	15.5

Table 2. The amount in percentage of the replaced non-labeled compounds and their SIL-IS in the blank matrix with their ^{13}C and/or ^{12}C fractions expressed by the ratio between the area of the non-labeled metabolite or SIL-IS and the sum of the areas of all isotopes of the particular compound detected in the blank matrix multiplied by 100.

Compound	Replaced amount in %
ATP	98
[$^{13}C_{10}$]ATP	98
ADP	98
[$^{15}N_5$]ADP	86
AMP	98
[$^{15}N_5$]AMP	87
Ac-CoA	100
[$^{13}C_2$]Ac-CoA	100
Mal-CoA	100
[$^{13}C_3$]Mal-CoA	100
NAD^+	99
NADH	100
$NADP^+$	100
NADPH	100

Table 3. Linearity, correlation coefficient (R^2), intra day (n=3) and inter day (n=9) precision, accuracy/recovery and LOD of the method for the measured compounds.

Compound	Linearity range (µg/ml)	R^2	Levels (µg/ml)	Intra day		Inter day		LOD (ng)
				Precision (RSD in %)	Accuracy (%)	Precision (RSD in %)	Accuracy (%)	
			175	11	96	7	99	
ATP	2.5-250	0.994	25	6	87	5	90	1.2
			7.5	3	98	6	105	
			70	10	97	6	100	
ADP	1-100	0.997	10	1	100	8	93	0.5
			3	3	100	12	100	
			70	8	110	10	106	
AMP	1-100	0.997	10	10	101	13	99	0.3
			3	12	107	14	94	
			70	5	95	8	102	
Mal-CoA	1-100	0.999	10	6	87	11	97	0.0004
			3	3	96	5	101	
			70	3	95	4	97	
Ac-CoA	1-100	0.998	10	6	93	10	103	0.002
			3	6	91	12	92	
			70	6	99	10	97	
NADH	1-100	0.993	10	6	96	18	98	0.001
			3	18	97	19	83	
			70	3	94	4	97	
NADPH	1-100	0.99	10	4	117	11	119	0.0005
			3	18	95	22	87	
			49	13	97	7	98	
$NADP^+$	0.7-70	0.992	7	9	111	14	113	0.0001
			2.1	1	83	8	95	
			49	10	98	6	99	
NAD^+	0.7-70	0.989	7	4	99	11	97	0.0004
			2.1	1	83	13	96	

Table 4. Matrix effects evaluated using the Matuzewski et al. [10] approach.

Compound	Slope of the linear regression		Matrix effects (%)[a]
	Standards in neat solution	Standards spiked in the matrix before the analysis	
ATP	141932	90326	64
ADP	175368	18573	11
AMP	197760	17878	9
Mal-CoA	125069	65668	53
Ac-CoA	100890	46985	47
NAD^+	413125	191866	46
NADH	103824	50004	48
$NADP^+$	905022	317781	35
NADPH	66962	25448	38

[a]Matrix effects are expressed as the ratio of the slopes of the calibration curves (1/x weighting) constructed from standards spiked before the analysis in matrix processed through the sample procedure to the standards prepared in the neat solution multiplied by 100. 100% indicates no matrix effects.

Table 5. Amount of oxidized species detected in the samples spiked with NADPH and NADH at three different concentrations and processed through the sample preparation procedure.

Compound	Spiked amount nmol/mgDW	Oxidized species nmol/mgDW (SD)[a]
NADH	0.12	0.1 (0.04)
	0.40	0.3 (0.06)
	2.81	2.3 (0.15)
NADPH	0.11	0.03 (0.002)
	0.36	0.1 (0.016)
	2.51	0.9 (0.155)

[a]SD – standard deviation; number of replicates per day: 3; number of days: 3

Table 6. Intracellular concentrations of the compounds of interest in *S. cerevisiae* expressed as nmol/mg dry weight measured using the developed ion-pair UHPLC-MS/MS method.

Compound	nmol/mg dry weight (SD)
AMP	0.32 (0.03)
ADP	0.91 (0.07)
ATP	9.13 (0.6)
Mal-CoA	0.03 (0.002)
Ac-CoA	0.78 (0.05)
NAD^+	6.24 (0.002)
NADH	0.12 (0.02)
$NADP^+$	0.71 (0.06)
NADPH	0.07 (0.01)

[a] (SD) standard deviation, n=3

Figure legends

Figure 1. Chromatograms of the labelled and non-labelled compounds spiked in the matrix before the extraction.

Figure 2. Measured (dashed line) and calculated (bars) isotopic patter of ATP extracted from *S. cerevisiae* cultivated in medium containing 50 % (w/w) [$^{13}C_6$]glucose/non-labelled glucose.

Figure 3. Superimposed chromatograms of [$^{15}N_5$]ADP, [$^{15}N_5$]AMP, ATP and Mal-CoA in the blank matrix (dashed line) and matrix spiked (solid line) with the corresponding standards with concentrations as given in the figure.

Figure 4. Overall process recovery of the analyzed compounds. Data shown are average of duplicate samples with their standard deviations.

Figure 5. Overlaid chromatograms of post column infusion of 10 µg/ml ATP into the MS after injecting eluent A or matrix blank with superimposed chromatograms of all the analyzed compounds.

6.6 Paper 6 - Multitargeted analysis of intracellular metabolites in various microorganisms using ion-pair reversed phase UHPLC-Q-TOF MS

Olivera Magdenoska, Subir Kumar Nandy, Anna Eliasson Lantz, Jette Thykær, Kristian Fog Nielsen.

Manuscript in preparation. It will be submitted to Analytical Biochemistry

Multitargeted analysis of intracellular metabolites from various microorganisms using ion-pair reversed phase UHPLC-Q-TOF MS

Olivera Magdenoska[1], Subir Kumar Nandy[1], Anna Eliasson Lantz[2], Jette Thykær[2], Kristian Fog Nielsen[1]

[1]Eucariotic Biotechnology, Department of Systems Biology, Technical University of Denmark, Søltofts Plads 221, DK-2800 Kgs. Lyngby, Denmark.

[2]Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søltofts Plads 228, DK-2800 Kgs. Lyngby, Denmark.

[3]Novo Nordisk A/S Hagedornsvej 1, DK-2820 Gentofte, Denmark

[*]Corresponding author. Tel: +45 45252725; Fax: + 45 45884148; E-mail: olima@bio.dtu.dk

Abstracts

Analysis of intracellular metabolites in microorganisms is important for understanding the regulation of the metabolic processes within the cells. Targeted LC-MS/MS are the methods of choice for analysis of these metabolites. These methods measure only targeted metabolites, thereby failing to detect major interference such as contaminants that has been introduced during the sample preparation or are coming from the sample itself. Here we investigate the use of QTOF accurate mass screening combined with auto MS/MS for measurement of intracellular metabolites extracted from *Streptomyces coelicolor*, *Microbispora corallina* and *Saccharomyces cerevisiae*. The analysis resulted in identification of 60 metabolites of which 40 were available as reference standards while the other 20 were tentatively identified based on retention time comparison to reference standards, MS/MS fragmentation, and accurate mass. Compared to the QqQ instrument sensitivity and specificity was improved for some compounds where specific low abundance daughter ions were needed by the QqQ instrument. The QTOF full scan data identified several major sources of ion-

suppressing contaminants, and also showed that the targeted QqQ detection did not overlook any major peaks.

Introduction

Intracellular metabolome consists of a highly complex and dynamic suite of molecules that interconnect and act as central players in many biochemical pathways within the cell [1,2]. The abundance of these metabolites can vary within the cell since their levels can change within seconds as a response to the changes in the environment [3-8]. Therefore the sample preparation for these metabolites consists of a quenching that will stop the metabolic activities followed by extraction [10-13]. The chemical properties of the metabolites in the extracts are widely diverse, however majority of them have masses below 1000 Da and are charged in aqueous solution [14]. A combination of different analytical platforms such as liquid chromatography–mass spectrometry (LC–MS), gas chromatography–mass spectrometry (GC–MS), or nuclear magnetic resonance spectrometry (NMR) are required for analysis of the whole suite of metabolites [13]. An alternative to this will be an analytical technique that will offer a broad coverage of metabolites involved in different pathways thus assisting in better understanding of the biochemical processes.

LC-MS is a widely applied technique for analysis of intracellular metabolites. In combination with an effective sample preparation, LC-MS provides a high specificity, sensitivity as well as wide dynamic range and can be utilized for analysis of polar to semi-nonpolar compounds. Triple quadrupole instruments are traditionally applied for analysis of intracellular metabolites. Although they offer high sensitivity and wide dynamic range [15], the number of metabolites that can be analyzed in a single run is limited, although increasing dramatically with developments in electronics lowering the time used per MRM [15]. None the less with high resolution instruments such as quadrupole time-of-flight (Q-TOF) and Orbi-trap, data can be collected across a wide mass range without decrease in sensitivity and due to its high

resolving power, it can distinguish between contaminants and signals of interest [15]. Therefore LC-HRMS is an effective technique for analysis of multiple compounds in complex biological matrices. Beside HRMS data, Q-TOF and Orbi-trap instruments offer two types of MS/MS acquisition, data depended and data independent. The advantage of the data dependent acquisition is that HRMS and MS/HRMS data are acquired in a single run without the need for precursor ion specification. However, not all ions will be fragmented, even using exclusion lists and time. The data independent acquisition (also called MS^E) is an alternative where all the ions generated will be fragmented using alternating low and high collision energies. Together with the accurate mass measurements this aid to the identification of unknown compounds.

The data acquired by high resolution instruments are more complex and their analysis can become quite laborious. Hundreds to thousands of peaks can be detected in a single run that might correspond to compounds of interest, contaminators, artefacts etc. Therefore, tools that can help in fast characterization and assignment of the chromatographic peaks have been in focus from the instrument manufactures. Several manufactures now handle targeted search lists on 1000-3000 compounds, where they for each compound search for one four characteristic ions (including fragments) and their accurate masses, chromatographic behavior, isotopic patterns. For the TOF instruments the software can also compensate for possible oversaturated scans in the peak apex where their mass accuracy is compromised [17]. Our group has for secondary metabolites refined the concept by building and improving the compound databases, organized in chemical database format as ACD for possible substructure searching and selection. These databases can then easily be converted to the MS vendor search formats.

The auto MS/MS data acquisition provides and extra dimension for fast and very accurate identification. With the standardized fragmentation energies used for acquiring the data, database can be used across research fields.

In this study a targeted ion-pair UHPLC-QqQ method is compared to a full scan accurate mass (QTOF). The methodology was used in combination with different quenching methods to evaluate not only how many compounds that were miss by the QqQ instruments but also to identify major interfering matrix components. By using the full scan and auto-MS/MS compounds not available as reference standards could be tentatively identified based on retention time, accurate masse, isotopic pattern and especially MS fragmentation.

Materials and methods

Chemicals

The standards were purchased from either Sigma Aldrich (Steinheim, Germany) or Santa Cruz biotechnology. [$^{13}C_{10}$]ATP and [$^{15}N_5$]ADP were purchased from Sigma–Aldrich (Steinheim, Germany) while [$^{15}N_5$]AMP was from Silantes (Munich, Germany). [$^{13}C_3$]Mal-CoA and [$^{13}C_2$]Ac-CoA were from Euriso-top (Gif-sur-Yvette Cedex, Paris, France). Tributylamine (TBA) (puriss plus grade), methanol (LC-MS and HPLC grade), CH_3Cl, 1,4-Piperazinediethanesulfonic acid (PIPES), EDTA and acetic acid (LC–MS grade) were obtained from Sigma–Aldrich. A Milli-Q system (Millipore, Bedford, MA, USA) was used for water purification.

Cell growth and sample preparation

Saccharomyces cerevisiae. Biomass was obtained from glucose chemostat cultures of *S. cerevisiae* (CEN.PK113-11C strain, dilution rate of 0.1 h^{-1}, biomass dry weight of 4.07g/l). Details of the cultivation have been previously described in Tomas et al. 2015. Samples were taken during steady state conditions in the end of continuous cultivation and were quenched and extracted as previously described by Villas-Boas et al. In brief, 5 mL of culture broth was

sprayed into pre-cooled (-40°C) falcon tube containing 20 mL of 60% methanol, spun down for 2 min at 5000xg in precooled centrifuge (-10°C) and extracted using boiling ethanol method [26] followed by evaporation under nitrogen. The samples were redissolved in 300 ul eluent A (10mM tributylamine and 10 mM acetic acid).

Microbisporra corralina samples was obtained from glucose chemostat cultures with dilution rate of 0.01 h^{-1}, nitrate as nitrogen source, pH 7.3, temperature 30 °C, biomass dry weight of 7.07 g/l, 2.5 ml of culture was taken from the fermentor and released into 10 ml of 60% (v/v) methanol/water solution, precooled to -40 °C in an ethanol/dry ice bath. The mixture of culture and quenching solution was immediately removed from the bath and centrifuged for 3 min at 4248 x g and 4 °C. After decanting, methanol/ CH_3Cl (1:2, v/v) mixture precooled in the ethanol/dry ice bath was added to the cell pellet followed by vortexing, followed by adding 2 ml of ice-cold buffer (PIPES, 3 mM; EDTA, 3 mM; pH 7.2). The spin tubes with the mixture were placed on a platform shaker for 60 min at 4 °C and 160 rpm, followed by centrifugation for 3 min at 4248 x g and 4 °C. The upper aqueous phase (3.7 ml) was collected and transferred to another tube. Two milliliters of methanol and 2 ml of ice cold buffer was added to the CH_3Cl phase followed by vortexing and centrifugation for 3 min at 4248g and 4 °C. The upper aqueous phase was collected (4 ml) and pooled with the previous extract. The extract was lyophilized and redissolved in 200 ul of eluent A (10mM tributylamine and 10 mM acetic acid).

*Streptomyces coelicolor s*amples were obtained from a batch cultivation using a minimal medium as described in Borodina et al. (2008). The extraction and quenching procedure are the same as for *Microbisporra corralina*. The only difference is that here 5 ml of culture was taken from the fermentor and quenched with 20 ml -40 °C 60% (v/v) methanol/water solution.

Ion pair chromatography

LC separation was performed on an Agilent 1290 UHPLC system (Agilent Technologies, Torrance, CA, USA) a previously described in Magdenoska et al. with 10 mM TBA as ion pair reagent. Eluent A: 10 mM TBA and 10 mM acetic acid, while eluent B was 90% MeOH (v/v) containing 10 mM TBA and 10 mM acetic acid.

Two different gradients were used: 19.5 min (0-12 min from 0 to 50 %B, 12-12.5 min from 50 to 100% B, 12.5 -14 min 100% B, 14-14.5 min from 100 to 0% B, 14.5-19.5 0% B) and 36 min (0-5 min 0% B, 5-10 min from 0 to 2% B, 10-11 min from 2 to 9% B, 11-16 min 9% B, 16-24 min from 9 to 50% B, 24-28.5 min 100% B, 28.5-30 min 100% B, 30-30.5 min from 100 to 0% B, and 30.5 to 36 min 0% B). The injection volume was 0.5 µl unless otherwise stated.

Q-TOF LC/MS

Agilent 6550 iFunnel QTOF LC/MS system equipped with a Dual Agilent Jet Stream ESI source was operated in negative ion mode and 2-GHz extended dynamic range at a resolution of 25,000 full width at half-maximum (FWHM). The QTOF instrument was tuned for fragile molecules by decreasing the potentials in the ion path by 5 V relative to the auto-tune values. The ion source parameters were as follows: gas temperature, 290 °C; gas flow 14 l/min; sheath gas temperature, 400 °C; sheath gas flow 11 l/min; nebulizer pressure, 45 psi; capillary voltage 3000 V; nozzle voltage 2000 V. Mass to charge ratio range of 50-1700 was used for acquiring both MS and MS/MS spectra with a scan rate of 3 spectra/sec. Auto MS/MS spectra were acquired using three distinct collision energies: 10, 20 and 40 eV. MS/MS spectra were acquired using an intensity cut-off of 10 000 counts, ion-exclusion time was 0.15 min. Exclusion list was created with the lock masses and the most abundant background ions, thereby increasing the fraction of metabolites that were fragmented. For the compound for which fragmentation pattern was not acquired by the auto MS/MS, targeted

MS/MS was conducted. Mass calibration was conducted by continuous pumping of a reference mass solution using isocratic pump at a flow of 1000 µl/min and a splitter (1:100) [15]. The reference mass solution contained 800 ml MeOH, 199 ml water, 1 ml of 1.1 mg/ml hexakis(2,2,3,3-tetrafluoropropoxy)phosphazine (Apollo Scientific Ltd., Cheshire, UK), and 200 µl of 10% TFA (v/v). The reference masses used for mass calibration were m/z 1033.9881 $[M+TFA-H]^-$ and m/z 112.9855 $[TFA-H]^-$.

Data mining of QTOF data for known unknowns

For identification of known unknowns (known metabolites not available as reference standards), negative ion ESI accurate mass data were obtained from the extracts using MS and auto MS/MS acquisition. Aggressive dereplication was done as described in Kildegaard et al. 2014 by creating *.csv files that contained the empirical formula and the name of the compounds listed in an in-house database created in the ACD chemfolder format (ref). This database was created by importing the online available metabolom databases of *Escherichia coli* (*E. coli* metabolome database, http://www.ecmdb.ca/) and yeast (Yeast Metabolome Database, http://www.ymdb.ca/) in the sdf format and adding all in-house available reference standards [Klitgaard et al. 2014]. The identification of the known unknowns was done using the find by formula feature of the Mass Hunter Qualitative analysis software (version B.06.00) using a mass tolerance of $\pm$ 10 ppm. The adducts ($[M-H]^-$ and $[M+CH_3COO]^-$) and dimers ($[2M-H]^-$) were included in the search and all ions were treated as being singly charged. A minimum score of 70 was used to ensure that only compounds with fitting isotope patterns were marked.

In addition, method, containing all identified compounds, was created in Mass Hunter Quantitative analysis software that facilitated the screening of different matrices for specific compounds as previously explained in Nielsen et al.2015. The MS/HRMS spectra acquired

were compared against the online available METLIN database. For compounds for which no fragmentation pattern was available in the METLIN database manual inspection of the spectra was done for loss of a phosphate, fragmentation of a C-O, C-N bonds.

Results and discussion

Chromatography

Prolongation of the gradient from 19 to 36 min was necessary for separating a number of compounds with the same elemental composition, especially the sugar phosphates (e.g. G6P, F6P, F1P, G1P). As their MS/MS spectra displayed the same ions albeit in slightly different abundances, their chromatographic separation was vital for quantification. In the case of the NTPs (e.g. ATP and dGTP have same composition, Table 1, specific fragments that correspond to the aromatic part of the molecule aided the identification when Auto-MS/MS was triggered.

When installing a new column in the UHPLC, overnight equilibration with the ion-pair eluent was necessary in order to obtain reproducible retention times, narrow peaks and stable run to run RT, showing a slow equilibration of the reagent to the silica backbone of the column. Variation (30-60 sec) of retention times were mainly observed for the mono-phosphorylated compounds eluting in the middle of the gradient, and was correlated to replacement of the eluents. When the same eluent was used within a sequence no changes in the RT was observed. The late eluting compounds were not affected with the change of the eluent batch.

The separation of some isomers was strongly affected by the column performance, e.g. could AMP and dGMP be baseline separated on a new column, but after 100-500 biological samples these two compounds co-eluted and had 30 % broader peaks and more pronounced tailing. Presumable due to fouling by polar compounds as high temperature and isopropanol

wash could not reconstitute the column. The separation of the sugar phosphates and the AMP and dGMP isomers was shown to be a good indicator for column replacement.

Sample preparation can also affected the LC peak shape of the compounds and is illustrated by the analysis of *M. corallina* extracts prepared using two different sample preparation methods (Figure 1).

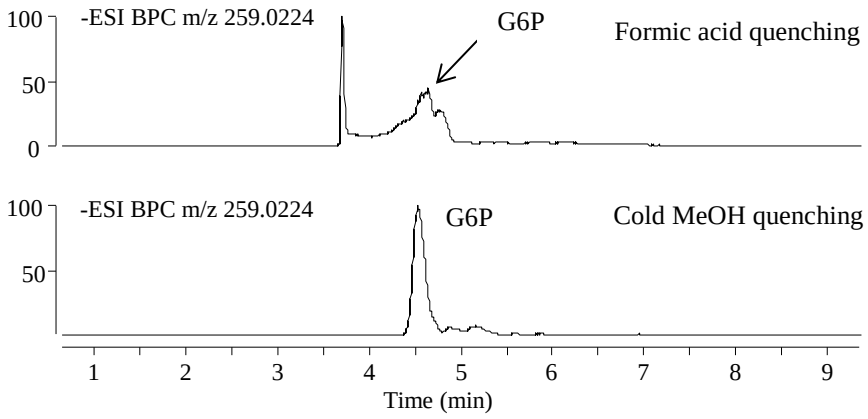


Figure 1. Effect of the sample preparation on the peak shape of glucose 6-phosphate extracted from *M. corralina*. The upper chromatogram is obtained from a sample quenched with formic acid and extracted by 3x freeze/thaw. The lower chromatogram is obtained from a sample quenched with 60% MeOH -40°C and extracted with MeOH/chloroform (2:1, v/v), showing the same RT as a pure reference standard.

It can be clearly seen that the peak shape of G6P (formic acid quenching) is altered as a result of the matrix compounds impairing the chromatography. This effect is more pronounced at the beginning of the gradient as also know from reversed chromatography where injection of sample in stronger solvent than the start gradient shows the same effects.

When a mixture of all the proteinogenic amino acids was analyzed, as expected the basic amino acids, lysine and arginine, were not retained at all (ion exclusion) and were eluted in the void volume of the column. The acidic amino acids were more retained and had an RT of 2.8 min for glutamic acid (~0.8 min) and 2.9 for aspartic acid showing some ion exchange characteristics. The aromatic amino acids, tyrosine (RT 1.3 min) and phenylalanine (RT 2.2

min) were also retained although their overall charge was zero (Figure 2). This could be explained by the reversed phase interactions between the aromatic part of the amino acids and the phenyl hexyl groups from the stationary phase. This shows that a combination of an-ion-pairing and reverse phase interactions are responsible for the retention of the compounds during IP-RP. Furthermore the tyrosine was eluting earlier due to the presence of the polar hydroxyl group.

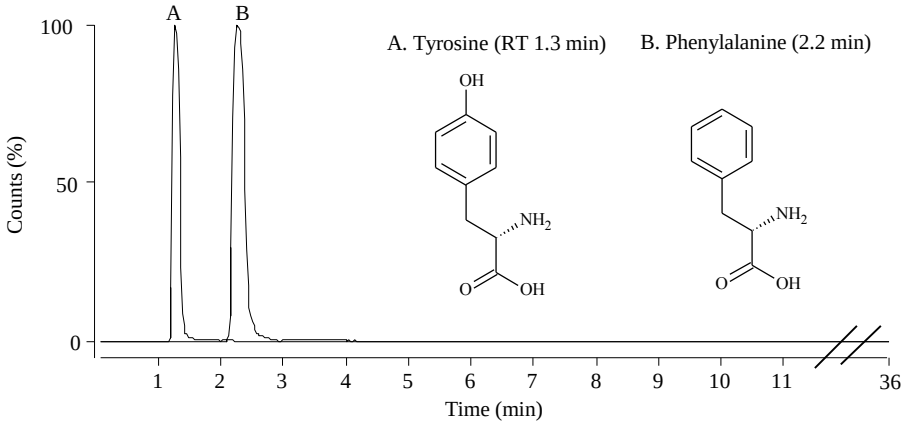


Figure 2. Chromatogram of tyrosine (RT 1.3min) and phenylalanine (RT 2.2 min) obtained by analyzing a standard mixture of amino acids using ion-pair LC-Q-TOFMS. Column used: Poroshell 120 Phenyl-Hexyl. Eluent A:10 mM TBA and 10mM acetic acid, eluent B: 90% MeOH containing 10 mM TBA and 10mM acetic acid. Gradient: 0-5min 0% B, 5-10 min 0-2% B, 10-11 min 2-9% B, 11-16 min 9% B, 16-24min 9-50% B, 24-28min 50%B, 28-28.5min 100%B, 28.5-30min 100%B, 30-30.5 100-0 %B, 30.5-36min 0%B.

3.2 QTOF-MS performance

Data on how the MS signal is affected by the injection solvent in which the standard solutions are prepared, were obtained by analyzing 0.5 µl of 50 µg/ml NADH, NADPH, 3-phosphoglycerate and succinic acid prepared in water or eluent A (Figure 3).

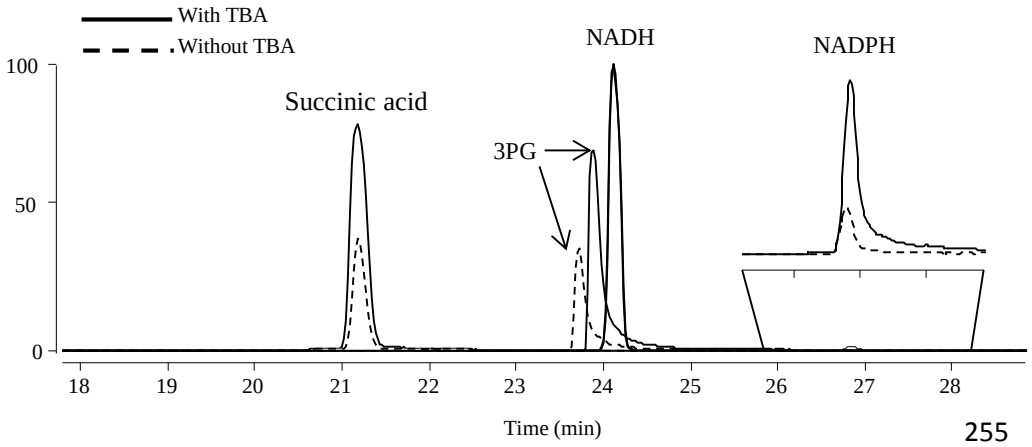


Figure 3. Effect of the injection solvent on the LC-MS signal intensity of succinic acid, 3PG, NADH and NADPH.

As shown in Figure 3, the LC-MS peak area of succinic acid, 3PG and NADPH was increased by addition of TBA in the injection vial, but this was not the case for NADH. In addition, slight change in the retention time was observed for 3PG when TBA was added.

3.3 Analysis of reference standard compounds

In order to inspect the predominant molecular ions, the adduct and dimer formations as well as the ion-source fragmentation, different classes of compounds ranging from sugar phosphates, nucleotides, coenzymes, carboxylic acids, amino acids and sugar phosphates were analyzed (Table 1).

For all the compounds, the analysis showed formation of $[M-H]^-$ and $[2M-H]^-$ ions, with $[M-H]^-$ as the predominant ion (Table 1). Acetate adducts were observed only for some of the coenzymes and the amino acids. Due to their very low intensity (below 5000 counts) they were not taken into consideration.

When using the find by formula algorithm for teh long lists of tentative compounds, double peak was assigned to erythrose 4-phosphate (E4P) with one apex at a RT of 7.3 min. and the other one at 8.4 min. The predominant ion of the first peak (RT 7.3 min.) was m/z 259.0224

which corresponded to the acetate adduct of E4P, while the second most intense ion was m/z 199.0013 that corresponded to $[M-H]^-$ of E4P. The peak with an RT of 8.4 min showed the opposite. However, extracting the EIC's of m/z 259.0224 and m/z 199.0013 both with mass tolerance of ±10 ppm resulted in two distinct peaks, which showed that the m/z 259.0224 detected at RT of 8.4 min. was due to the tailing of the first peak (RT of 7.3). Since all the other sugar phosphates showed $[M-H]^-$ as the predominant ion and no acetate adducts, the peak with a RT of 8.4 was considered to be corresponding to E4P while the peak with lower RT to be as a result of a contamination probably with sugar phosphates that have m/z 259.0224 as previously reported by *Luo et al*. The MS/MS fragmentation pattern of the ions with m/z 259.0224 (RT of 7.3 min) and m/z 199.0013 (RT of 8.4 min.) were identical to the fragmentation patterns of the hexose phosphates and E4P in the METLIN database, respectively. Furthermore, when compared to the other sugar phosphates, the intensity of the $[M-H]^-$ of E4P was lower, probably due to the more pronounced ion source fragmentation that was observed in the full scan spectra.

In general loss of one phosphate group was observed for the phosphorylated compounds, loss of one or two CO_2 group for carboxylated compounds, while the amino acids showed loss of CO_2 or NH_3 (Table 1). The most pronounced ion-source fragmentation was observed for cis-aconitic acid where the area of the $[M-H]^-$ was around 50% of the area of both of the fragments. For all the other compounds the ion-source fragments were lower than the molecular ion. This could be avoided by changing the settings in the ion source, however it would result in decreased sensitivity for other compounds.

cGMP was formed by in-source fragmentation of c-diGMP. On the other hand the dimer of cGMP had the same mass of c-diGMP (m/z 689.0876). When a mix of cGMP and c-diGMP was analyzed, the EIC of m/z 689.0876 showed two chromatographic peaks one at a RT of cGMP and the other one at an RT of c-diGMP. It should be noted that conversion of c-

diGMP into cGMP did not occur in the vial which was confirmed by analyzing single pure standards of cGMP and cdiGMP that showed only one chromatographic peak for both cGMP (RT 20 min.) and cdiGMP (RT 23.9 min.).

Table 1. Molecular ions, adducts and dimers formation and ion source fragmentation of the in-house available standards of primary metabolites.

Compound	RT	$[M-H]^-$	$[2M-H]^-$	Ion source fragmentation
Lys	0.6	+	-	-
Arg	0.6	+	-	Loss of CN_2H_2
His	0.7	+	-	Loss of NH_3 or $-CO_2$
Ser	0.8	+	-	-
Cys	0.8	+	-	-
Thr	0.8	+	-	Loss of C_2H_4O
Pro	0.9	+	-	-
Val	0.9	+	-	-
Met	1.1	+	-	-
Tyr	1.3	+	-	-
Phe	2.3	+	-	Loss of NH_3
Glu	2.9	+	-	Loss of H_2O or $-CO_2$
Asp	3.0	+	-	Loss of $-CO_2$
E4P	8.4	+	+	-
G6P	7.1	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
G3P	7.7	+	+	-
F6P	7.6	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
M6P	7.9	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
Gal6P	8.4	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
G1P	8.8	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
X5P	9.8	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
Ru5P	10.3	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
CMP	11.9	+	+	-
F1P	11.5	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
dCMP	13.4	+	+	-
DHAP	12.9	+	+	$[H_2PO_4]^-$, $[PO_3]^-$
UMP	14.2	+	+	-
NAD	14.3	+	+	-
GMP	16.5	+	+	-
IMP	16.8	+	+	-
dUMP	17.0	+	+	-
dGMP	19.2	+	+	-
dTMP	19.4	+	+	-
AMP	19.8	+	+	-
cGMP	20.0	+	+	c-diGMP
Succinate	20.3	+	-	Loss of H_2O and $-CO_2$

cAMP	21.5	+	+	-
UDP-Glucose	22.3	+	+	-
3PG	22.8	+	+	-
2PG	22.8	+	+	-
6PG	22.8	+	+	-
XMP	22.9	+	+	-
Fumarate	22.9	+	-	Loss of -CO_2
α-KG	22.4	+	-	Loss of -CO_2
NADH	23.1	+	+	NAD
UDP	23.4	+	+	UMP
GDP	23.4	+	-	GMP
PEP	23.6	+	+	PO_3^-
c-diGMP	23.9	+	+	cGMP
cis-Aconitic acid	23.9	+	-	Loss of -CO_2
ADP	23.9	+	+	-
NADP	23.9	+	+	-
Citric acid	24.0	+	-	-
GTP	25.7	+	+	GDP
dCTP	25.7	+	+	-
UTP	25.8	+	+	UDP
dGTP	25.9	+	+	-
ATP	26.1	+	+	ADP
dTTP	26.2	+	+	-
dATP	26.5	+	+	-
NADPH	26.8	+	+	NADP
FAD	27.1	+	+	-
Acetyl-CoA	29.5	+	+	-
Propionyl-CoA	29.5	+	-	-
Malonyl-CoA	29.5	+	-	-
Methylmalonyl-CoA	29.5	+	-	-

3.3 Finding the ion-suppressing peaks

3.4 Auto MS/MS

Collision energies of 10 and 20 eV were enough to fragment most of the molecules while CE of 40 eV was in general necessary to fragment larger stable molecules such as FAD or NADPH.

3.5 Aggressive dereplication combined with MS/HRMS library search

Using the search lists, more than 300 peaks were assigned to different compounds and by using the Mass Hunter Quantitative software, very fast inspection of the chromatograms was possible in order to eliminate compounds that were assigned to peaks that were practically noise or blanks.

The criteria used to confirm the tentative identification of the known unknown included: i) adduct formation, ii) RT and iii) MS/MS fragmentation pattern.

As described in the previous section, $[M-H]^-$ was detected as the predominant ion for the different classes of compound. Therefore when an acetate adduct was detected in the extracts as the predominant ion those results were taken with precaution. In general it was observed that the m/z of an acetate adduct of one compound was the same with the m/z of the $[M-H]^-$ of another compound, thereby resulting in false identification.

Based on the structural formula of the compound suggested as a possible match, a prediction of the retention time was possible. Due to the usage of an ion-pair chromatography with tributylamine as ion pair reagent an RT increase was expected when the overall negative charge of the molecules increased by increasing the number of present COO^- or PO_4^- groups and decreasing of the number of groups e.g. NH_2 that can be positively charged under the pH used during the analysis.

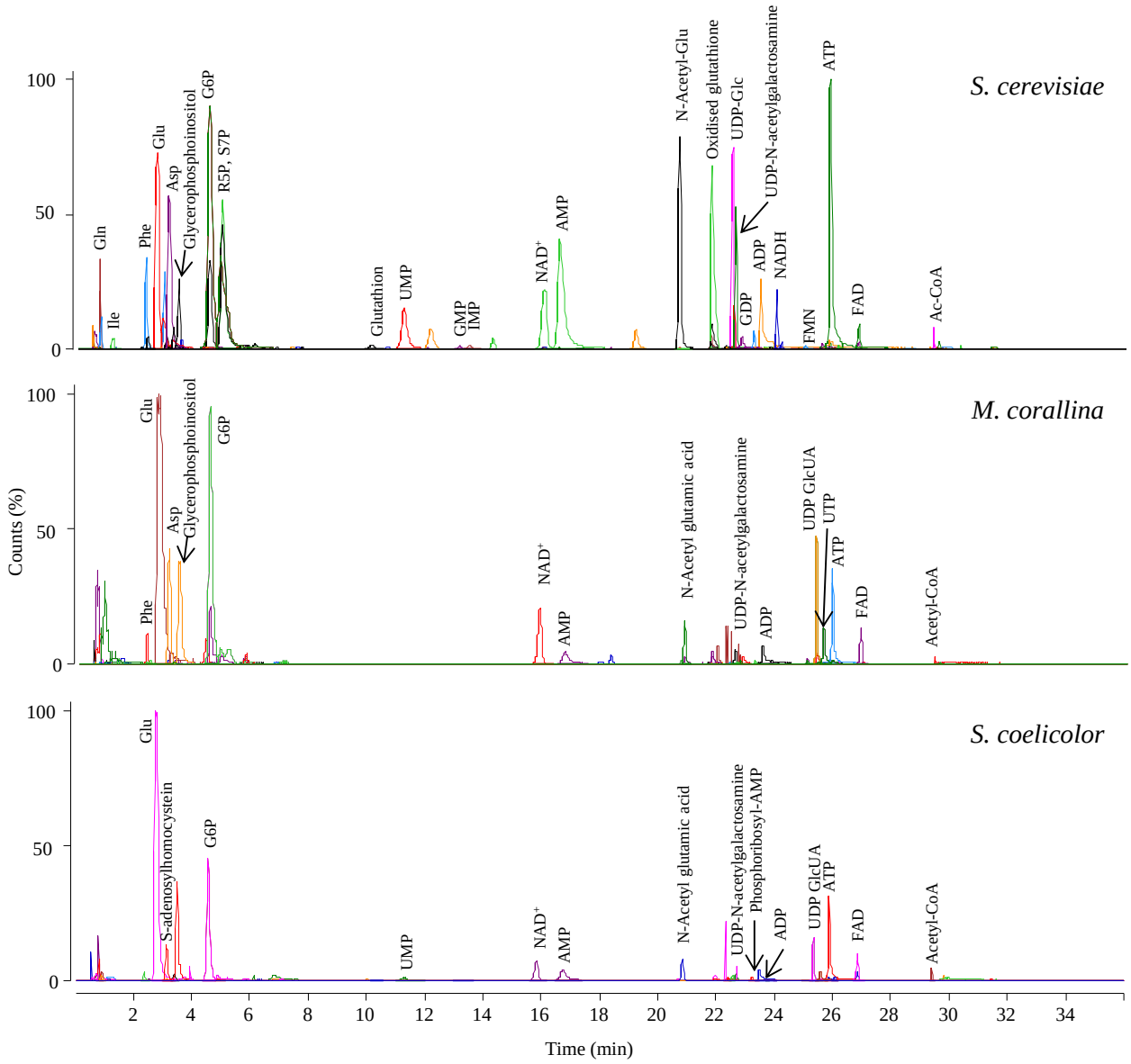
The MS/MS fragmentation patterns were used to confirm the putative identities that were assigned by the software to a particular chromatographic peak. The combination of RT and MS/HRMS spectra increased the confidence of the identification and was taken as the limiting factor for accepting the putative identities that were suggested by the software.

The search list combined with the Find by Formula Feature provided a quick overview of the possible isomers that might be responsible for a certain chromatographic peak. This was especially useful when chromatographic separation of the isomers was not possible.

ATP/dGTP, UDP-Glc/UDP-Gal are some of the many examples of isomers present in the cell extracts.

3.6 Identification of known unknowns

The untargeted approach using the aggressive dereplication combined with the METLIN library for searching MS/HRMS data was used to identify as many intracellular metabolites (known unknowns) as possible in three different matrices: *S. cerevisiae, M. corralina* and *S. coelicolor*. Figure 1 shows the highlighted peaks of the identified compounds from the three different extracts.



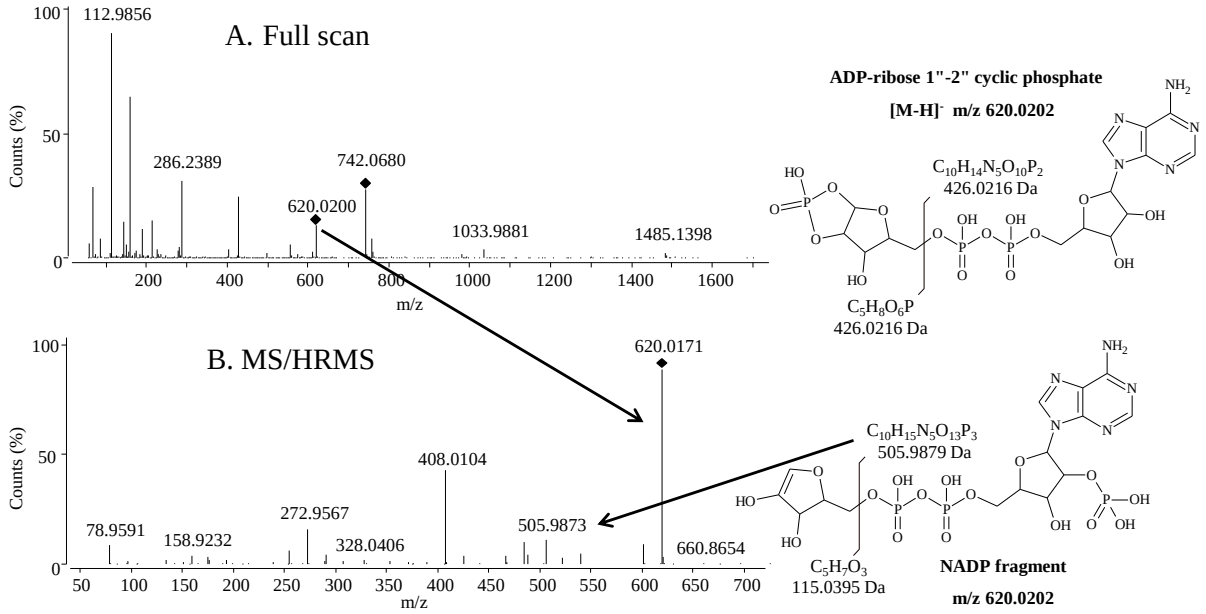
347 **Figure 4.** Base peak chromatograms *Saccharomyces cerevisiae*, *Microbisporra corralina* and
348 *Streptomyces coelicolor* extracts in negative ESI mode. Peaks of the compounds identified
349 using MS/HRMS are highlighted

350 Compounds, ranging from amino acids, sugar phosphates, nucleotides and organic acids were
351 identified using the aggressive dereplication approach (**Table 2**).

352 **TABLE 2 [In progress]**

353 Although the data analysis approach described in this study yielded putative identifications
354 for several hundreds of compounds for a very short time, one should also be aware of the

pitfalls of this approach. One example of misidentification due to an in-source fragmentation using this approach was the detection of the sugar nucleotide ADP-ribose 1″-2″ cyclic phosphate. The aggressive dereplication approach led to tentative identification of ADP-ribose 1″-2″ cyclic phosphate (m/z 620.0202) at a RT of 23.15 min. This was a compound for which a standard in-house was not available. Furthermore by using the aggressive dereplication approach NADP (m/z 742.0680) was also identified and had the same RT as ADP-ribose 1″-2″ cyclic phosphate. From previous measurements it was found that an in-source fragmentation of NADP gives a fragment ion with m/z 620.0202 $[C_{15}H_{21}N_5O_{16}P_3\text{-}H]^-$ which has the same m/z ratio as ADP-ribose 1″-2″ cyclic phosphate. This indicated a possible misidentification of ADP-ribose 1″-2″ cyclic phosphate due to the in-source formation of the fragment from NADP with the same m/z ratio as ADP-ribose 1″-2″ cyclic phosphate that further on gives positive match. When inspecting the fragmentation pattern of ADP-ribose 1″-2″ cyclic phosphate a fragment $[C_{10}H_{15}N_5O_{13}P3]^-$ (m/z 505.9836; 9 ppm) was detected that corresponds to a fragment that is specific for NADP, leading to the conclusion that ADP-ribose 1″-2″ cyclic phosphate was misidentified.



373 **Figure 5. A.** Full scan showing the [M-H]$^-$ of NADP (m/z 742.0680) and a fragment from
374 ions source fragmentation of NADP (m/z 620.0202); **B.** MS/HRMS of m/z 620.0202 showing
375 a fragment with an m/z 505.9873.

376 The predominant ion detected by the aggressive dereplication can also be an indicator
377 whether the identification result should be taken with a precaution. By using the aggressive
378 dereplication a peak with an RT of 11.2 min had two IDs, UMP and D-4-hydroxy-2-
379 oxoglutarate. The predominant ion for UMP was [M-H]$^-$ while for D-4-hydroxy-2-
380 oxoglutarate was [2M-H]$^-$, both with m/z of 323.0286. The acetate adduct and the [M-H]$^-$ of
381 D-4-hydroxy-2-oxoglutarate were also detected and had 5 times lower intensity when
382 compared to the [2M-H]$^-$ ion of the same compound. As described before the predominant
383 ions for the standards analyzed were the [M-H]$^-$ ions. Furthermore, when compared to other
384 carboxylic acids such as fumaric and succinic acid, the RT of 11.2 min was too low for an
385 organic acid with two COOH groups. Furthermore, the measured MS/MS spectra of m/z
386 323.0286 was identical to the MS/MS of UMP and did not contain any of the usual fragments
387 for carboxylated compound such as loss of CO_2 groups. Therefore it was concluded that the
388 peak detected at a RT of 11.2 min. corresponded to UMP. Furthermore the ions that

corresponded to $[M+CH_3COO]^-$ and $[M-H]^-$ of d-4-Hydroxy-2-oxoglutarate, were found to be a background ions that were present over the whole chromatogram.

When using the aggressive dereplication approach very often more than one peak is assigned to one compound. In this case, the RT criteria together with the MS/HRMS data can help solving this problem. An example for this is the identification of glutathione which was only detected in *S. cerevisiae* extract. Two chromatographic peaks were assigned to glutathione when using the aggressive dereplication approach. An acetate adduct of the glutathione m/z 366.0972 (accuracy of 1.3 ppm) was detected at RT of 3.06 min while $[M-H]^-$ m/z 306.0760 (accuracy 1.6 ppm) was detected at 10.08 min. An MS/HRMS data was acquired from both the acetate adduct $[M+CH_3COO]^-$ and the $[M-H]^-$. MS/HRMS data in the library were only available for glutathione but not for his acetate adduct. When the extracted MS/HRMS spectra of the $[M+CH3COO]^-$ were compared with the glutathione fragmentation pattern in the METLIN library, no similarity was found. This leaded to the conclusion that the peak at a RT of 3.06 min is not glutathione. Furthermore the extracted MS/HRMS spectra of the $[M-H]^-$ at a RT of 10.08 min. were compared to the glutathione fragmentation spectra in the METLIN library which resulted in a positive match. In addition oxidized (RT 21.7 min) and dicarboxyethyl glutathione (RT 23.9 min) were also detected.

…more cases of positive or misidentification

Conclusion

Besides the tested on 60 reference standards further 20 compounds could be tentatively identified based on retention time comparison to reference standards, MS/MS fragmentation, and accurate mass. For method development the QTOF was superior for evaluating the reasons for ion-suppression and differences in sample preparation.

Compared to the QqQ instrument sensitivity and specificity was improved for some compounds (e.g. ATP, AMP) where specific low abundance daughter ions were needed by the QqQ instrument. The QTOF data allowed to verify that the compounds detected were not fragments from ion source fragmentation of co-eluting compounds. Most importantly the full scan QTOF data allows retrospective data analysis for both new compounds as well as decomposition products.

References

[1] E. Nevoigt, Progress in Metabolic Engineering of *Saccharomyces cerevisiae,* Microbiol. Mol. Biol. Rev. 72 (2008) 379-412.

[2] J. L. Adrio, A. L Demain, Recombinant organisms for production of industrial products, Bioeng. Bugs. 2 (2010) 116–131.

[3] S. Dietmair, N. E. Timmins, P. P. Gray, L. K. Nielsen, J. O. Krömer, Towards quantitative metabolomics of mammalian cells: Development of a metabolite extraction protocol, Anal. Biochem. 404 (2010) 155–164.

[4] S.G. Villas-Bôas, U. Roessner, M.A.E. Hansen, J. Smedsgaard, J. Nielsen (Eds.), Metabolome Analysis: An introduction, Wiley–Interscience, Hoboken, NJ, 2007, pp. 39–82.

[5] W. de Koning, K. van Dam, A method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH, Anal. Biochem. 204 (1992) 118–123.

[6] S.G. Villas-Bôas, J. Højer-Pedersen, M. Åkesson, J. Smedsgaard, J. Nielsen, Global metabolite analysis of yeasts: Evaluation of sample preparation methods, Yeast 22 (2005) 1155–1169.

[7] W. M. van Gulik, Fast sampling for quantitative microbial metabolomics Curr. Opin. Biotechnol. 21 (2010) 27–34.

[8] M. Rizzi, M. Baltes, U. Theobald, M. Reuss, In vivo analysis of metabolic dynamics in *Saccharomyces cerevisiae*: II. Mathematical model, Biotechnol. Bioeng. 55 (1997) 592–608.

[9] S. G. Villas-Bôas, S. Mas, M. Åkesson, J. Smedsgaard, J. Nielsen, Mass spectrometry in metabolome analysis, Mass Spectrom. Rev. 24 (2005) 613-646

[10] D. Visser, G. A. van Zuylen, J. C. van Dam, M. R. Eman, A. Pröll, C. Ras, L. Wu, W. M. van Gulik, J. J. Heijnen, Analysis of in vivo kinetics of glycolysis in aerobic *Saccharomyces cerevisiae* by application of glucose and ethanol pulses, Biotechnol. Bioeng. 88 (2004) 157-167.

[11] A. B. Canelas, W. M van Gulik, J. J. Heijnen (2008) Determination of the cytosolic free NAD/NADH ratio in *Saccharomyces cerevisiae* under steady-state and highly dynamic conditions. Biotechnol. Bioeng. 100 (2008) 734–743.

[12] V.M. Boer, C. A. Crutchfield, P. H. Bradley, D. Botstein, J. D. Rabinowitz, Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations, Mol. Biol. Cell 21 (2010) 198-211.

[13] S. G. Villas-Bôas, S. Mas, M. Åkesson, J. Smedsgaard, J. Nielsen, Mass spectrometry in metabolome analysis, Mass Spectrom. Rev. 24 (2005) 613-646.

[14] R. M. Seifar, C. Ras, J. C. van Dama, W. M. van Gulik, J. J. Heijnen, W. A. van Winden, Simultaneous quantification of free nucleotides in complex biological samples using ion pair reversed phase liquid chromatography isotope dilution tandem mass spectrometry, Anal. Biochem. 388 (2009) 213–219.

[15] S. Bajad, V. Shulaev, Highly-parallel metabolomics approaches using LC-MS2 for pharmaceutical and environmental analysis, Trends Anal. Chem. 26 (2007) 625–636.

[16] W. Lu, B. D. Bennett, J. D. Rabinowitz, Analytical strategies for LC–MS-based targeted metabolomics, J. Chromatogr. B 871 (2008) 236–242.

[17] A. Klitgaard, A. Iversen, M. R. Andersen, T. O. Larsen, J. C. Frisvad, K. F. Nielsen, Aggressive dereplication using UHPLC-DAD-QTOF: screening extracts for up to 3000 fungal secondary metabolites, Anal. Bioanal. Chem. 406 (2014)1933-1943.

6.7 Paper 7 (non-peer reviewed)- Metabolomets ioniske komponenter bestemt ved kromatografi og massespektrometri

Olivera Magdenoska, Daniel Killerup Svenssen, Peter Boldsen Knudsen, Andrea Thorhallsdottir, Mhairi Workman and Kristian Fog Nielsen.

Published in Dansk Kemi 2015

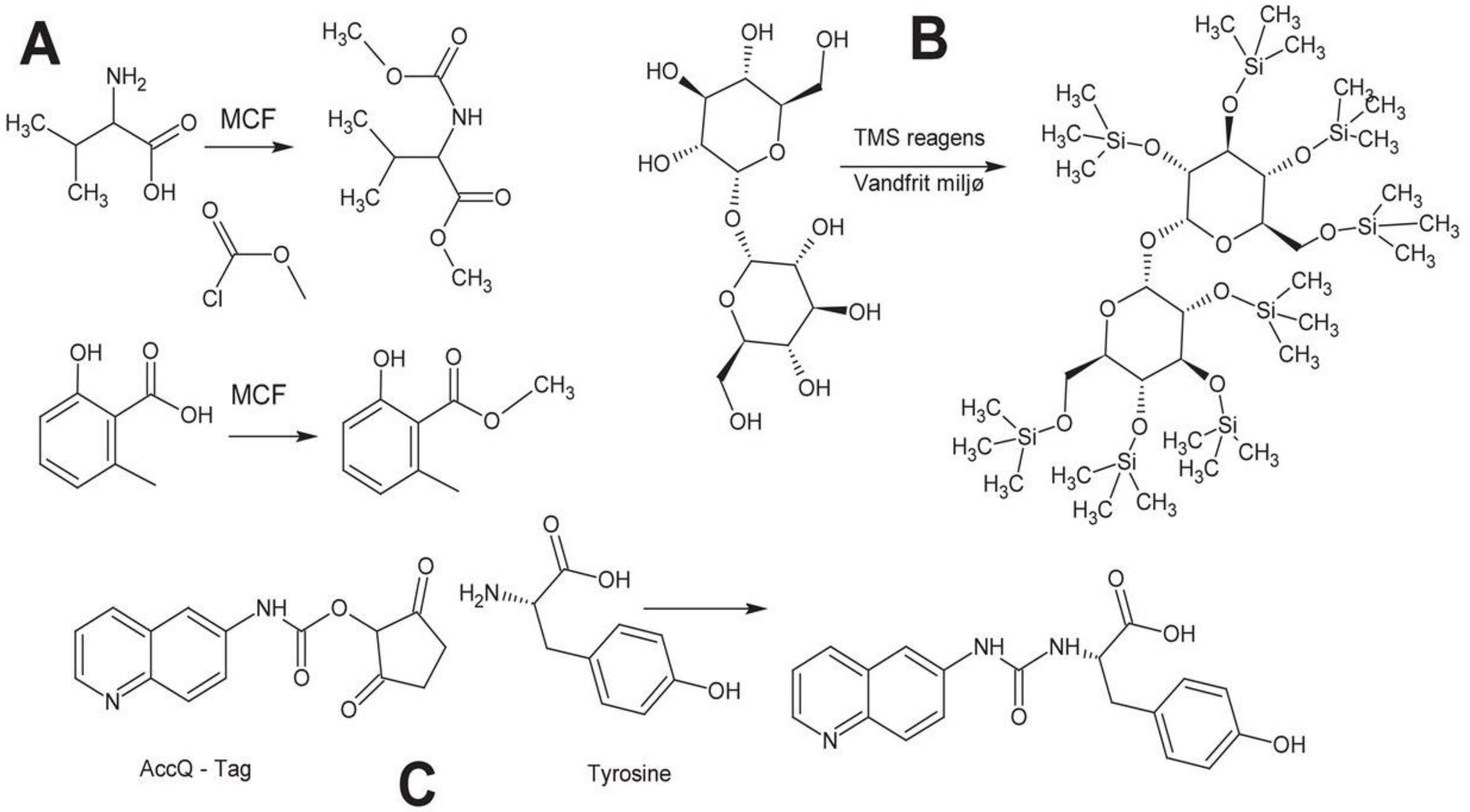
Metabolomets ioniske komponenter bestemt ved kromatografi og massespektrometri

Bestemmelse af pool-størrelser af fosforylerede intracelllulære energi- og redoxkomponenter, samt cellens byggesten er vigtig for at afgøre, om metabolismen i cellefabrikker er påvirket af høj produktionsbelastning. Det er afgørende, at kunne evaluere, hvordan forskellige genetiske modifikationer påvirker det intracellulære maskineri.

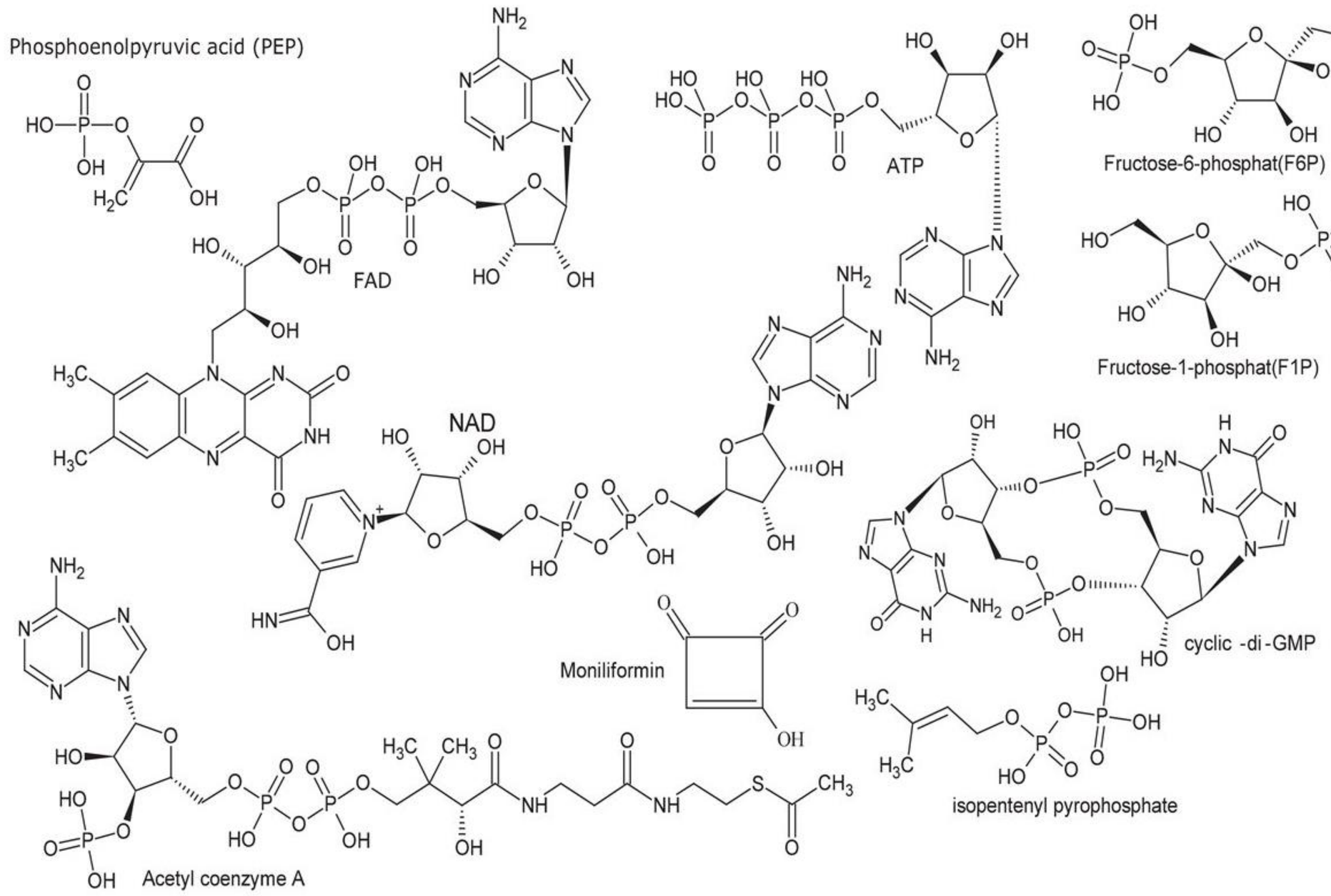
Af Olivera Magdenoska[1], Daniel Killerup Svenssen[1], Peter Boldsen Knudsen[1], Andrea Thorhallsdottir[2], Mhairi Workman[1] og Kristian Fog Nielsen[1]
[1]Institut for Systembiologi, DTU
[2]Actavis, Island

Analyse af de mange intracellulære metabolitter er en udfordrende opgave. Den kan udføres ved en kombination af kromatografi og massespektrometri (MS), mens bestemmelse af hele det intracellulære metabolom, kræver en kombination af flere analysemetoder.

Aminosyrer og andre organiske syrer kan med fordel analyseres med gaskromatografi-massespektrometri (GC-MS), som derivater med f.eks. chloroformat eller trimethylsilyl (TMS), figur 1 A og B. Aminosyrer og aminer kan derivatiseres med OPA eller AQC, figur 1 C, og efterfølgende analyseres med HPLC-UV eller mere følsomme metoder så som HPLC-fluorescens eller HPLC-MS.



Figur 1. Derivatisering af aminosyre og phenolisk syre med chloroformat (A), silylering af sukker (B), AccQ-Tag derivatisering af aminosyre (C).



Figur 2. Eksempler på forskellige nøglemetabolitter i cellestofskiftet samt mykotoksinet, moniliformin.

Hovedparten af energimetabolitter (ATP, ADP, GTP etc.), redox co-faktorerne (FAD, NADH, NADP etc.), sukkerfosfater, byggesten til polyketider (acetyl-CoA, malonyl-CoA, etc.), og terpener (isopentenyl og dimethylallyl fosfater), samt byggestene til DNA og RNA, er alle fosforylerede, figur 2. Da GC-MS-analyse ikke er anvendelig til fosforylerede metabolitter (enkelte mono-fosforylerede stoffer kan bestemmes som TMS-derivater), efterlades et kæmpe hul i metabolomanalysen.

Til separation af fosforylerede forbindelser anvendes kapillarelektroforese eller væskekromatografi før MS-detektion. Vi vælger at satse på en selektion af HPLC/UHPLC-teknikker, da vi har erfaring inden for dette område.

Den primære udfordring med de fosforylerede stoffer er, at de ikke tilbageholdes ved omvendt fase-kromatografi, der ellers er den mest effektive metode mht. peak-kapaciteten og separation af isomerer. Det gælder også, hvis man bruger de mere polære omvendte fase-materialer som phenyl, pentafluorophenyl og biphenyl.

Ionbytning og hydrofil interaktionskromatografi (HILIC) udgør de eneste reelle alternativer til omvendt fase-kromatografi. Begge findes i mange varianter.

I første omgang satsede vi på HILIC, da denne teknik udnyttes til svampetoksinet, moniliformin [1] (pKa 0.5, figur 2). HILIC, vandig normalfase kromatografi udnytter det vandlag, der dannes over en polær stationær fase, samt dipol-dipol og ioniske interaktioner med den stationære fase. HILIC-faser fås i mange varianter, fra relativt svage interaktioner som:
i) diol- og amid-faser
ii) til intermediat interaktionskolonner som silica og silica hydrid [2]
iii) ioniske HILIC-faser som amino-silica og den zwitter-ioniske Zic-Hilic [3].

Specielt amino-silica er attraktivt, da det er muligt at manipulere forholdet mellem NH_2- og NH_3^+-grupper på overfladen ved at ændre pH.

Ingen af disse kolonner resulterede i stabile retentionstider, og for en række vigtige analytter var toppene for brede (op til seks min.). Behovet for 50 mM acetat til eluering af de trifosphorylerede nucelotider var desuden problematisk, da det krævede hyppig rensning af elektrospraykilderne på vores daværende Waters LC-MS-instrumenter.

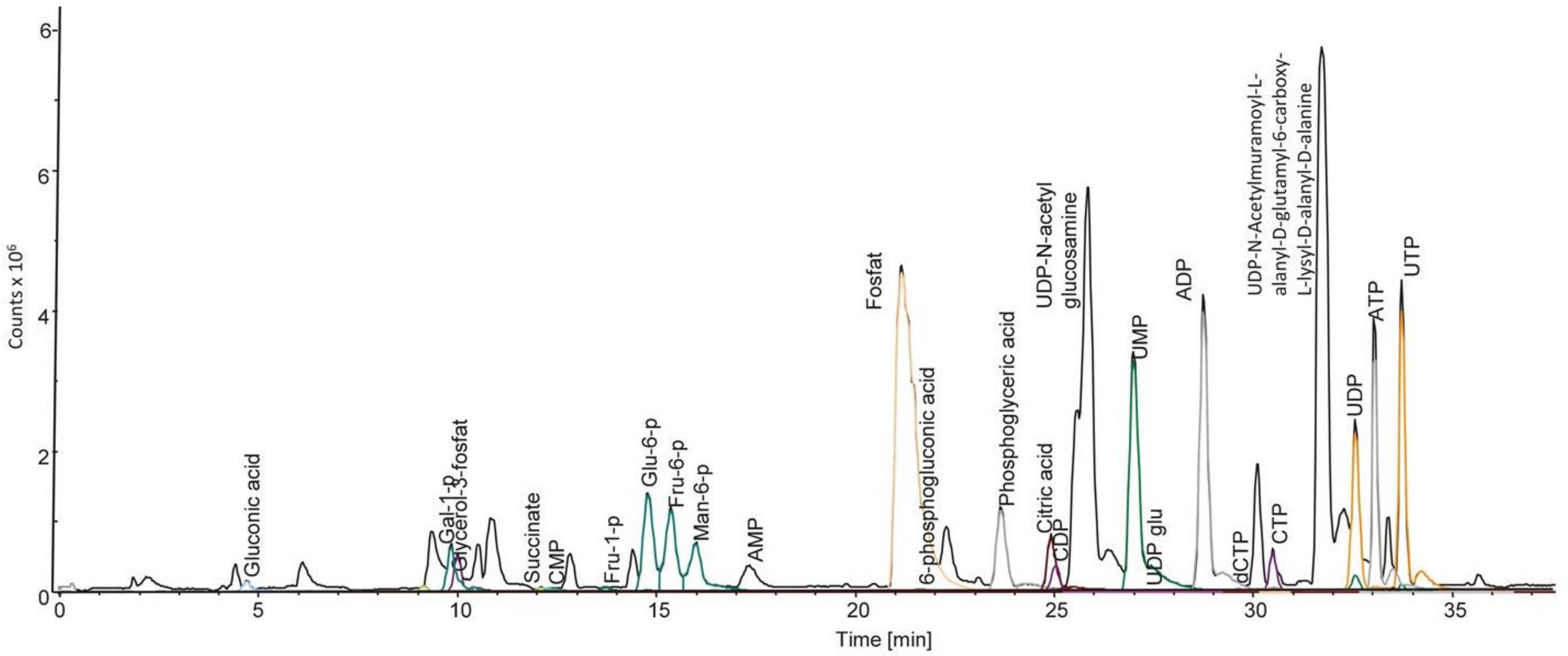
Dertil har HILIC-kolonnerne et begrænset dynamisk område, og det var nødvendigt at anvende ca. 90% acetonitril for at opnå god retention for organiske syrer samt polære ikke-ioniske stoffer. Desværre blev top-bredderne under disse betingelser endnu værre for de di- og tri-fosphorylerede stoffer, muligvis pga. udfældning i injektionssystemet.

Klassisk ionbytning med saltgradienter er udelukket, da saltene inden for få minutter blokerer. Derimod er høj-pH ionbytning, også kaldet ionkromatografi (IC) en mulighed, når en suppressor anvendes til fjernelse af K^+ eller Na^+ fra den gradient (typisk 10-100 mM) af NaOH eller KOH, der benyttes til separationen. Metoden er anvendelig til analyse af sukkerfosfater og mange di- og trifosphorylerede nucleotider, figur 3, side 10, og med en anden kolonne også sukre og oligosakkarider.

Metoden har den ulempe, at høj-pH labile stoffer, som de vigtige redox co-faktorer, acetyl-, malonyl-CoA'er hydrolyseres, så de ikke kan differentieres. En række andre metabolitter er ikke høj pH-stabile, og vi vælger derfor at satse på ion-par kromatografi til intracellulære metabolitter. IC-MS anvendes dog til sukre samt enkelte fosforylerede og sulfonerede stoffer.

Ion-par kromatografi

Det sidste alternativ er ion-par kromatografi, figur 4, side 10, ▶



Figur 3. IC-MS-analyse af *Streptomyces lividans* ekstrakt med (Dionex IC2100 med IonPac®AS11-HC), 1-100 mM NaOH gradient, kombineret med en Bruker maXis G3 Time of Flight MS med ESI- kilde.

en slags mobil ionbytter, der opbygges af ion-par modifieren, som er et ionisk og hydrofobt stof, der absorberes på overfladen af en omvendt-fase kolonne, typisk ved at blande 5-20 mM af modifieren i eluenten. Herved ændres mængden af ionbytter-grupper på kolonnen under kromatografien. Det er ikke tilrådeligt at ændre den mere end 10-30%, da dette resulterer i ikke reproducerbare retetionstider. Ved analyse af negativt ladede stoffer bruges en positivt ladet ion-par-danner og tilsvarende til positivt ladede stoffer, figur 4.

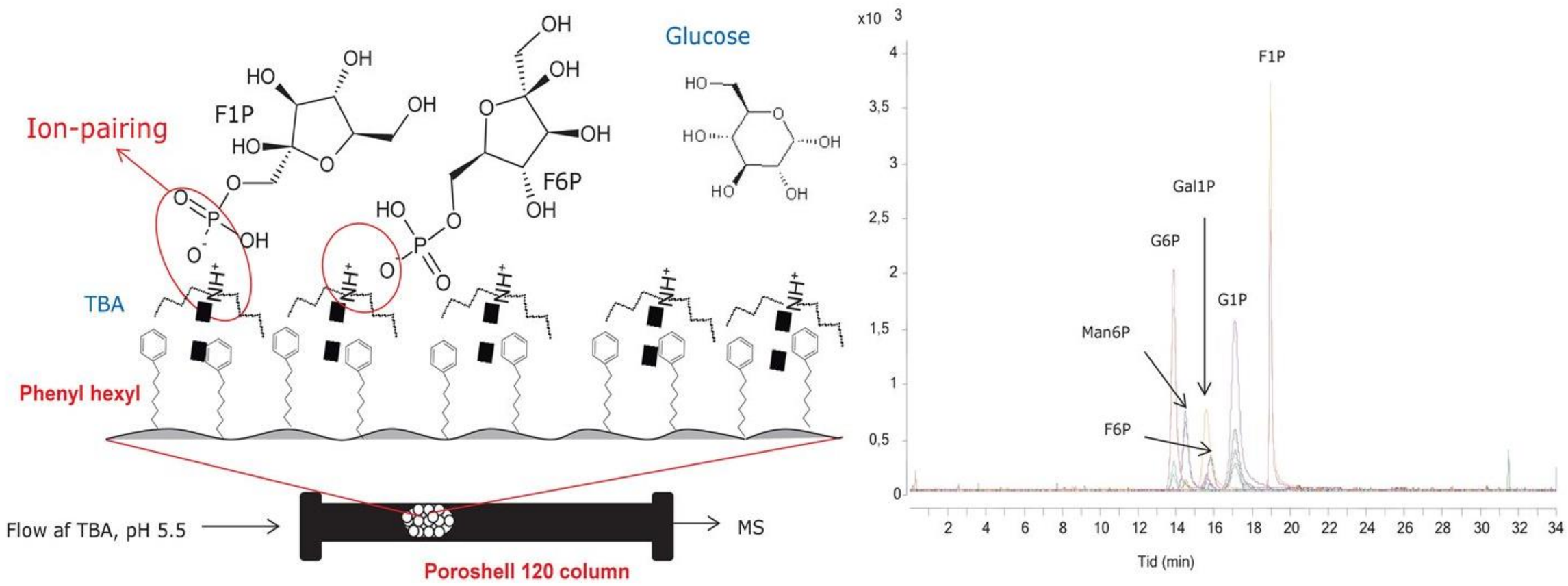
Hvis ion-par kromatografi skal bruges sammen med MS detektion, skal to krav opfyldes:

i) Ion-par-reagenset skal være flygtigt, så ion-kilden ikke blokeres, og

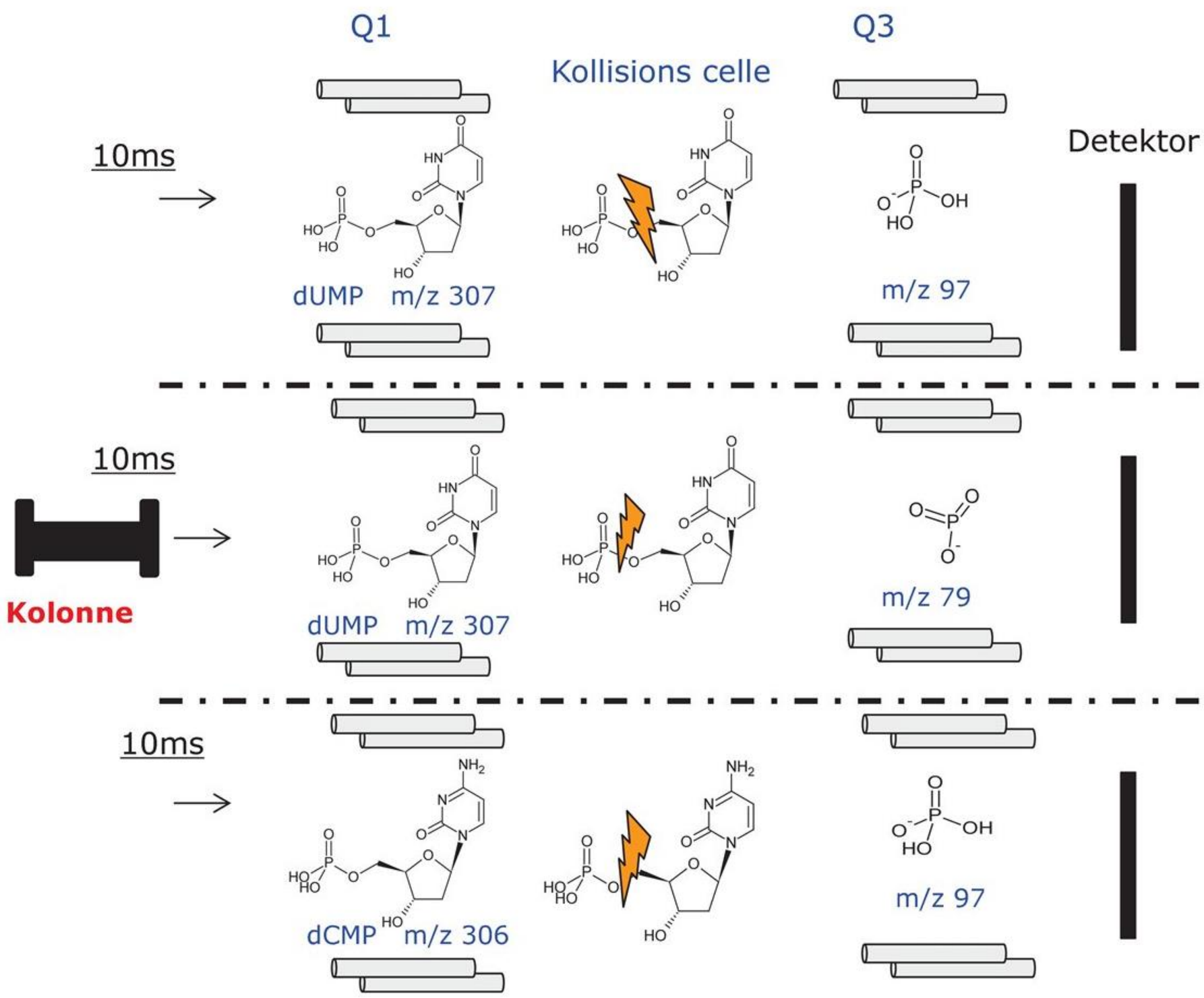
ii) MS'en skal opereres i den modsatte polaritet af reagensets ladning. Da alle de intracellulære metabolitter har en negativ ladning ved neutralt pH, anvendes tributylamin, figur 4, der kun tillader brug af MS-instrumentet i negativ polaritet. De positivt ladede ion-par reagenter er næsten umulige at fjerne fra HPLC-systemet igen, og et dedikeret system er nødvendigt til disse analyser.

Metoden resulterer i god kromatografisk selektivitet og tillader bl.a. separation af de fleste sukkerfosfater, der kemisk set er meget ens, figur 4. Detektionsmæssigt bruger vi primært tandem massespektrometri (MS/MS) på en tripel quadrupol MS, figur 5, med negativ electrospray (ESI) ionisering. Instrumentet er så hurtigt, at flere stoffer i praksis kan måles på én gang. Dette skyldes den korte fragmenteringstid (ca. 5-50 ms), der typisk udføres to gange pr. stof. Med 20 sekunders kromatografisk topbredde kan man således måle 50 stoffer, hvis der foretages 20 målinger hen over toppen. Da isotopmærkede interne standarder anvendes, er der i praksis brug for tre fragmenteringer (overgange) pr. stof, hvorfor vi af følsomhedshensyn bruger 30 ms pr. fragmentering.

MS/MS er pt. den mest følsomme metode, men den lider af den skavank, at kun kendte stoffer med kendte fragmenteringsmønstre kan detekteres. Derfor sammenkobles ion-par UHPLC-systemet til en quadrupol-Time-of-Flight MS med jævne mellemrum. Denne tilbyder mere følsomhed i full-scan mode og tillader akkurat masse MS. Det betyder, at ATP kan måles med høj nøjagtighed, *m/z* 505.98859874±0.001 [M-H],



Figur 4. Ion-par kromatografi med tributylamin som ion-par-modifier, non-anioniske hydrofile analytter som glucose tilbageholdes ikke. Til højre ses separationen af en række sukkerfosfater (G glukose, F fruktose, Gal galaktose, Man mannose).



Figur 5. Tandem MS-detektion af to stoffer på en tripel quadrupol MS, pga. af den lave måletid pr. stof kan mange stoffer måles på én gang.

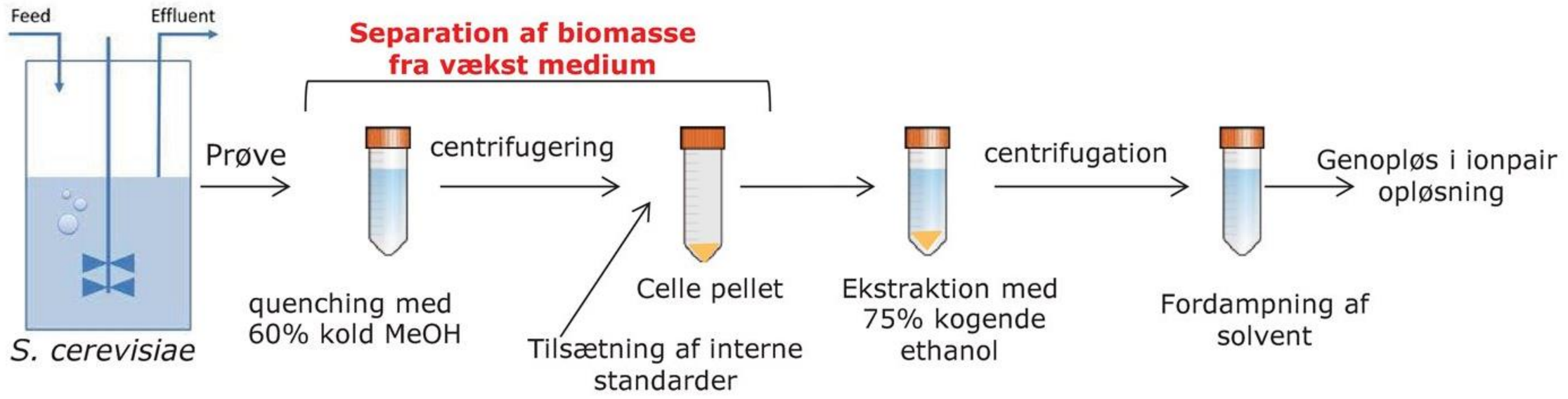
hvilket betyder, at elementarsammensætningen for kendte stoffer kan verificeres, selv uden en referencestandard. Instrumentet kan også anvendes til MS/MS, hvorved en identifikation af et kendt stof uden referencestandard, kan underbygges.

Quenching og prøveforberedelse

Som beskrevet er separation og detektion af de intracellulære metabolitter en yderst vanskelig analyseopgave, der kompliceres yderligere af en tidskrævende og besværlig prøveforberedelse. Grundet en meget høj turn-over rate af intracellulære metabolitter er det nødvendigt at stoppe metabolismen på under 1 sekund, således at de intracellulære pools ikke ændres signifikant; en proces der betegnes quenching. For at kunne evaluere effektiviteten af quenchingen anvendes energy charge ratio,

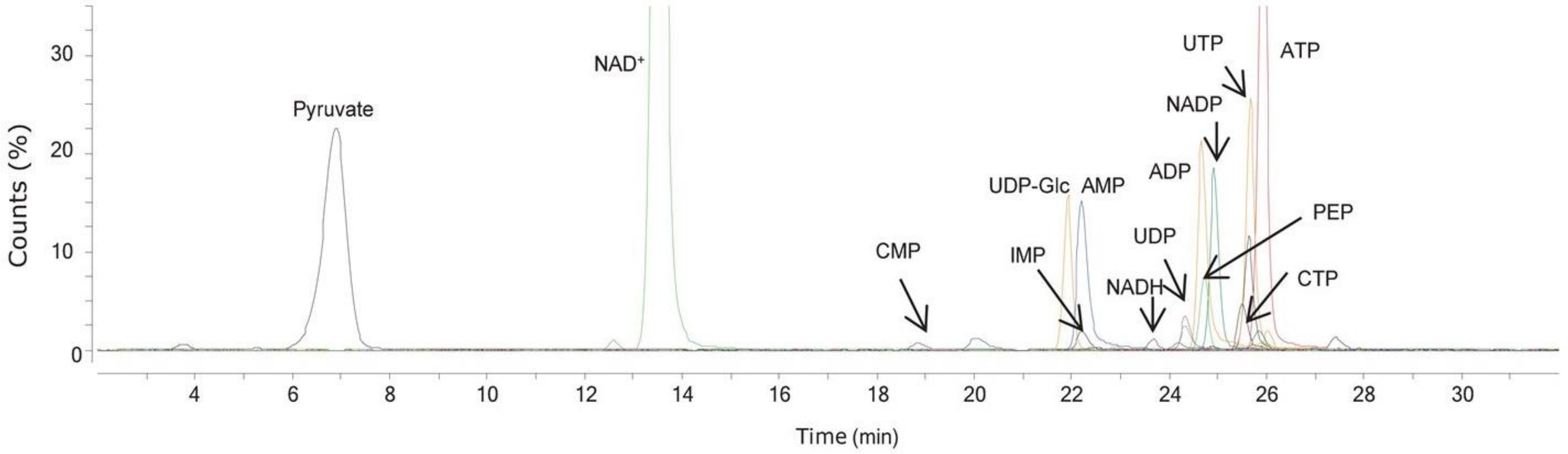
$$E = (C_{ATP} + 0.5 \times C_{ADP}) / (C_{ATP} + C_{ADP} + C_{AMP}),$$

hvor C er koncentrationen. Ratioen bør typisk ligge på 0,9-0,95, da normale celler indeholder væsentligt mere ATP end ADP og AMP. Der er tale om en særdeles sensitiv balance, som kræver høj præcision under prøveforberedelsen. Selv små unøjagtigheder kan forskubbe den ned til 0.5. Typisk kræves del en øvelse, før nye folk har succes med proceduren. ▶



Figur 6. Quenching af bagegær og klargøring af prøve til UHPLC-MS/MS-analyse.

Da der ikke kendes en universel quenching-metode, der dækker samtlige organismetyper, er det yderst vigtigt at optimere metoden til den valgte organisme, bl.a. for at sikre lav lækage og en høj energy charge ratio.



Figur 7. Ion-par UHPLC-MS/MS-analyse af et ekstrakt fra bagegær. Alle analytterne kan differentieres via deres forskellige MS/MS-fragmenteringer.

Den nemmeste metode er den som anvendes til gær, der kan quenches i 60% methanol ved -40˚C, uden at cellemembranen ødelægges. Så centrifugeres cellerne ned, og medie og methanol fjernes, hvorefter biomassen ekstraheres i kogende ethanol eller methanol-chloroform. Metoden sikrer, at lipider og en række interfererende stoffer fjernes via chloroformfasen. Den resterende vand-methanol-fase frysetørres, genopløses og analyseres, figur 7.

Andre celler såsom mammale er for sensitive til, at denne metode kan anvendes, da cellemembranen nedbrydes af methanol. Den bedste metode til denne type celler er quenching i iskoldt saltvand efterfulgt af forsigtig centrifugering, således cellerne ikke lyseres. Denne metode er til gengæld ikke anvendelig til gær og andre mikroorganismer, da disse indeholder enzymer, der er aktive selv ved lavere temperaturer.

Mælkesyrebakterier udgør en særlig svær gruppe at quenche, da der ikke findes en pålidelig metode, der sikrer membranstabiliteten. Det betyder, at de intracellulære metabolitter må analyseres sammen med interfererende materiale fra vækstmediet, resulterende i en kompleks blanding med store mængder sukker og andre medieelementer, i essens en sirup. Vi vælger derfor at oprense prøver på aktivt kul [4] for at kunne måle nucleotiderne, hvilket betyder tab af non-aromatic stoffer som f.eks. co-enzymer og sukkerfosfater.

Den sidste større gruppering af organismer inkluderer de filamentiøse bakterier og skimmelsvampe, der grundet den filamentiøse natur er særlig sensitive og let nedbrydes. Til denne type mikroorganismer anvendes 60% methanol ved -40˚C efterfulgt af enten centrifugering eller filtrering. Centrifugering er den foretrukne metode, da den er mere skånsom end filtrering, der kun anvendes i de tilfælde, hvor det ikke er muligt at bundfælde cellerne. Biomassen ekstraheres herefter med methanol-chloroform, hvorpå metabolitterne ekstraheres i en vandig buffer, der frysetørres. Som standard accepteres en lækageprocent på 20-30%.

Perspektiver

I øjeblikket arbejder vi med forskellige *proof of concept*-projekter, hvor molekylærbiologer har udviklet forskellige højproducerende skimmelsvampestammer til f.eks. produktion af 6-methylsalisylsyre, vanilliner, orsellinsyrer og mycophenolsyre med en produktion på op til 2 g/L og udbytter op til 10% på C-mol basis.

Skimmelsvampe kan dog producere en lang række uønskede stoffer, der besværliggør oprensningen af det ønskede produkt. Samtidig har denne type mikroorganismer ofte en kompleks morfologi, der besværliggør dyrkning i bioreaktorer samt en kompleks regulering, der ofte gør, at de ønskede stoffer kun produceres på faste medier. Det er derfor attraktivt at producere de ønskede stoffer i en mere simpel cellefabrik.

For at løse disse problemer og gøre processerne nemmere arbejder molekylærbiologer på at overføre de kodende gener til gær, som er generelt anerkendt som sikker og ikke producere uønskede toksiner. Samtidig er gær langt nemmere at dyrke i bioreaktorer, og har ofte højere vækshastigheder, hvorfor disse anses som ideelle til cellefabrikker.

Det har desværre vist sig at være svært at opnå samme udbytte i gær som i skimmelsvampe. En mulig forklaring er, at gær, som ikke naturligt producerer sekundære metabolitter, har en strømlinet metabolisme, der sikrer maksimal vækshastighed, hvorfor den ikke kan akkommodere højere produktion af ikke essentielle metabolitter.

Dette kommer til udtryk ved fermentering af en række gensplejsede gærstammer, hvor oscillerende vækst er observeret som funktion af højere kopiantal af det kodende gen. Vi arbejder med hypotesen, at denne oscillerende adfærd skyldes en redox ubalance, der kompenseres ved vekslen mellem to vækstformer.

Det er interessant at undersøge, hvordan *Aspergillus* til forskel fra gær adapterer sig, og hvordan de intracellulære pools ændres i forhold til produktion af det ønskede produkt. Sådanne informationer kan udpege flaskehalse i den centrale kulstofmetabolisme og derved guide molekylærbiologerne til yderligere optimering af cellefabrikken.

E-mail:
Kristian Fog Nielsen: kfn@bio.dtu.dk

Referencer

1. Sørensen JL, Nielsen KF, Thrane U. Analysis of moniliformin in maize plants using hydrophilic interaction chromatography. *J Agric Food Chem* 2007; 55:9764-9768.
2. Matyska MT, Pesek JJ, Duley J, Zamzami M, Fischer SM. Aqueous normal phase retention of nucleotides on silica hydride-based columns: Method development strategies for analytes revelant in clinical analysis. *J Sep Science* 2010; 33:930-938.
3. Lu W, Bennett BD, Rabinowitz JD. Analytical strategies for LC-MS-based targeted metabolomics. *J Chromatogr B* 2008; 871:236-242.
4. Magdenoska O, Martinussen J, Thykaer J, Nielsen KF. Dispersive solid phase extraction combined with ion-pair ultra high-performance liquid chromatography tandem mass spectrometry for quantification of nucleotides in Lactococcus lactis. *Analytical Biochemistry* 2013; 440:166-177.

Supplementary material

Table S1. Optimized precursor and product ion, fragmentor voltage, collision energy and cell accelerator voltage for the intracellular metabolites of interest during this Ph.D. study

Compound	Precursor Ion	Product Ion	Fragmentor Voltage	Collision energy	Cell Accelerator Voltage
2PG	185	97	90	20	3
HMG-CoA	910.1	425.9	135	45	4
HMG-CoA	910.1	407.9	135	50	4
3PG	185	97	90	15	3
6PG	274.9	96.9	100	10	4
Ac-CoA	808.1	426	100	30	4
Ac-CoA	808.2	408	100	40	4
ADP	425.9	158.9	120	25	3
ADP	425.9	134	115	25	3
AMP	346	211.1	100	15	3
AMP	346	79	90	30	3
ATP	505.9	407.9	100	20	3
ATP	505.9	273	95	30	3
cAMP	328.1	134	115	25	3
cAMP	328.1	78.9	120	35	3
c-diGMP	689	537.9	110	30	4
c-diGMP	689	149.9	100	35	4
CDP	401.9	158.9	130	25	3
CDP	401.9	79	90	40	3
cGMP	344	150	115	25	3
cGMP	344	132.9	135	35	3
Cis-aconitic acid	173	85	60	10	3
Citric acid	191	111	95	10	4
Citric acid	191	43	90	30	3
CMP	322	97	90	20	3
CMP	322	79	120	40	3
CTP	481.9	384	110	20	4
CTP	481.9	159	125	30	4
dATP	489.9	391.9	115	25	4
dATP	489.9	158.9	130	35	3
dCMP	306	195	130	15	3
dCMP	306	79	130	40	3
dCTP	465.9	367.9	100	20	4
dCTP	465.9	158.9	120	25	3
dGMP	346	195.1	100	20	3
dGMP	346	79	130	20	3

Table S1. Continued.

Compound	Precursor Ion	Product Ion	Fragmentor Voltage	Collision energy	Cell Accelerator Voltage
dGTP	505.9	257	115	35	3
dGTP	505.9	158.9	90	30	3
DHAP	169	78.7	70	35	3
DHAP	169	96.9	60	10	3
dTMP	321	194.9	125	15	3
dTMP	321	124.9	110	20	3
dTTP	480.9	383	90	20	4
dTTP	480.9	159	100	25	3
dUMP	307	195	110	15	3
dUMP	307	111	100	20	4
E4P	199	79	100	30	3
E4P	199	97	100	10	3
F1P	259	198.9	110	10	3
F1P	259	138.9	100	10	3
F1P	259	97	90	10	3
F1P	259	79	85	40	3
F6P	259	96.9	90	20	3
F6P	259	79	100	30	3
F6P	259	198.9	110	10	3
F6P	259	138.7	90	5	3
FAD	784.1	437	135	25	3
FAD	784.1	346.1	135	35	4
FBP	339	241	100	10	4
FBP	339	97	100	20	4
Fumaric acid	115	70.8	80	5	1
G1P	259	198.9	110	10	3
G1P	259	138.9	100	10	3
G1P	259	79	110	25	3
G1P	259	97	115	10	3
G3P	169	97	130	5	3
G3P	169	79	130	30	3
G6P	259	96.9	100	10	3
G6P	259	79	100	30	3
G6P	259	198.9	110	10	3
G6P	259	138.7	90	5	3
Gal1P	259	198.9	110	10	3
Gal1P	259	138.7	90	5	3
Gal1P	259	97	100	30	3
Gal1P	259	79	100	30	3

Table S1. Continued.

Compound	Precursor Ion	Product Ion	Fragmentor Voltage	Collision energy	Cell Accelerator Voltage
GDP	442	344.1	110	15	3
GDP	442	159	120	30	3
GMP	362	211	100	15	4
GMP	362	79	110	25	3
GTP	521.9	423.9	130	20	3
GTP	521.9	159	100	35	3
IMP	347	135	95	30	3
IMP	347	79	115	35	3
Malonyl CoA	852	807.9	135	25	4
Malonyl CoA	852	407.9	130	45	4
Man6P	259	198.9	110	10	3
Man6P	259	138.7	95	5	3
Man6P	259	96.8	90	15	3
Methylmalonyl CoA	866	822	135	25	3
Methylmalonyl CoA	866	407.9	115	40	3
NAD^+	662	540	120	15	4
NAD^+	662	407.9	120	35	4
NADH	664	397	110	30	4
NADH	664	346	110	35	4
$NADP^+$	742	620	110	15	3
$NADP^+$	742	408	110	35	3
NADPH	744	408	100	35	4
NADPH	744	397	100	35	4
OMP	367	97	100	25	3
OMP	367	79	95	25	3
Oxaloacetic acid	131	86.6	110	5	4
Oxaloacetic acid	131	43	110	5	4
PEP	167	97	70	20	3
PEP	167	79	70	10	3
PEP	167	97	70	20	3
PEP	167	79	70	10	3
ppGpp	601.9	503.9	130	24	1
ppGpp	601.9	423.9	130	24	3
Propionyl CoA	822	425.9	125	35	4
Propionyl CoA	822	407.9	120	40	4
Pyruvate	87	44	70	10	3
Pyruvate	87	43	70	5	3
Ribu5P	229	79	70	30	3
Ribu5P	229	97	60	15	3

Table S1. Continued.

Compound	Precursor Ion	Product Ion	Fragmentor Voltage	Collision energy	Cell Accelerator Voltage
Succinic acid	117	72.9	80	10	3
UDP	402.9	111	115	20	4
UDP	402.9	79	115	40	3
UDP-Glc	564.9	385	125	30	4
UDP-Glc	564.9	323	120	30	3
UMP	323	97	90	20	3
UMP	323	79	100	40	3
UTP	482.9	384.9	100	20	4
UTP	482.9	159	95	30	3
XMP	363	211.1	100	20	4
XMP	363	151.2	125	30	4
Xylu5P	229	79	30	3	3
Xylu5P	229	97	5	3	3
Xylu5P	229	138.8	5	3	3
ZMP	337	125	130	25	4
ZMP	337	97	125	15	4
ZMP	337	78.9	115	35	4
α-KG	145.1	57.3	80	10	5
α-KG	145.1	101.2	80	5	1

Table S2. Optimized precursor and product ion, fragmentor voltage, collision energy and cell accelerator voltage for the SIL-IS used during this Ph.D. study

SIL-IS	Precursor Ion	Product Ion	Fragmentor Voltage	Collision energy	Cell Accelerator Voltage
[^{13}C]GTP	531.9	433.9	115	20	3
[^{13}C]GTP	531.9	159	135	35	3
[^{13}C]ATP	515.9	417.9	125	20	3
[^{13}C]ATP	515.9	278	125	30	3
[^{13}C^{15}N]dGTP	521.1	262	100	35	3
[^{13}C^{15}N]dGTP	521.1	159.1	110	35	3
[^{15}N]dATP	495	396.7	115	20	4
[^{15}N]dATP	495	158.7	115	30	3
[^{15}N]UTP	485.1	387	90	20	3
[^{15}N]UTP	485.1	158.9	100	30	4
[^{13}C^{15}N]CTP	493.9	396	115	30	3
[^{13}C^{15}N]CTP	493.9	159	115	30	3
[^{13}C^{15}N]TTP	493	395	120	25	4
[^{13}C^{15}N]TTP	493	159	120	30	3
[^{13}C^{15}N]dCTP	477.9	379.9	125	20	4
[^{13}C^{15}N]dCTP	477.9	158.9	125	25	3
[^{13}C^{15}N]GMP	377.1	215.7	115	20	4
[^{13}C^{15}N]GMP	377.1	79.1	115	25	4
[^{15}N]AMP	350.8	211.1	100	15	4
[^{15}N]AMP	350.8	79	90	35	3
[^{13}C^{15}N]dGMP	361.1	200	100	20	4
[^{13}C^{15}N]dGMP	361.1	79	115	20	3
[^{13}C^{15}N]UMP	334.1	97	120	25	3
[^{13}C^{15}N]UMP	334.1	79	120	35	3
[^{13}C^{15}N]CMP	334	97	90	25	3
[^{13}C^{15}N]CMP	334	78.9	90	35	3
[^{13}C^{15}N]TMP	333	200	90	15	3
[^{13}C^{15}N]TMP	333	131.8	90	25	3
[^{13}C^{15}N]dCMP	318	200	90	15	4
[^{13}C^{15}N]dCMP	318	78.8	115	30	4
[$^{13}C_2$]Ac-CoA	810	462.8	135	35	3
[13C2]Ac-CoA	810	407.9	130	45	4

Table S2. Continued

[$^{13}C_3$]Mal CoA	855	810.2	110	30	4
[$^{13}C_3$]Mal CoA	855	463	110	25	4