Hyphenated mass spectrometric methods for quantitative metabolomics in *E.coli* and human cells

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

an der Fakultät für Chemie und Pharmazie

der Universität Regensburg

vorgelegt von

Birgit Timischl

aus Fürstenfeld, Österreich

Juli 2008

Diese Doktorarbeit entstand in der Zeit von April 2004 bis Juni 2008 am Genome Technology Center der Stanford Universität, Palo Alto, CA, USA, sowie am Institut für Funktionelle Genomik der Universität Regensburg.

Die Arbeit wurde angeleitet von Prof. Dr. Peter J. Oefner.

Promotionsgesuch eingereicht im Juli 2008

Kolloquiumstermin: 08.09.2008

Prüfungsausschuß: Vorsitzender: Prof. Dr. Manfred Liefländer

                                   Erstgutachter: Prof. Dr. Otto S. Wolfbeis

                                   Zweitgutachter: Prof. Dr. Peter J. Oefner

                                   Drittprüfer: Prof. Dr. Jörg Heilmann

1 Table of Contents

Danksagung

Ich freue mich sehr, nun endlich allen Leuten danken zu können, die auf unterschiedlichste Art und Weise zum Gelingen dieser Arbeit beigetragen haben.

An erster Stelle gilt mein Dank Prof. Peter Oefner für die Vergabe des interessanten Themas. Er lehrte mich was wissenschaftliches Arbeiten wirklich bedeutet und schärfte meinen Blick fürs Detail. Nur dadurch war es möglich, die erzielten Ergebnisse auf ein festes Fundament zu stellen.

Ein besonderes Dankeschön geht auch an Prof. Zlatko Trajanoski von der TU Graz, der mir den Beginn meiner Doktorarbeit ermöglicht und mich in der ersten Zeit finanziell unterstützt hat. Außerdem möchte ich mich bei Prof. Marina Kreutz und Prof. Petra Hoffmann von der Abteilung für Hämatologie am Klinikum Regensburg für die gute und fruchtbare Zusammenarbeit, sowie die interessante Einführung in das Gebiet der Immunologie bedanken.

Ganz herzlich möchte ich mich bei der Leiterin der Metabolomics-Gruppe Dr. Katja Dettmer bedanken, ihre Unterstützung und Ermutigung waren unendlich wertvoll, von ihr habe ich gelernt, was Analytische Chemie bedeutet. Ich werde die gemeinsamen Abende am Institut vermissen. Bei Hanne möchte ich mich besonders dafür bedanken, dass sie mir die Gewissheit gab, in allen Lebenslagen auf ihre Unterstützung zählen zu können. Beiden auch ein großes Dankeschön für die Durchführung aller GC-MS Messungen. Anne war die beste Bürokollegin, die man sich nur wünschen kann. Allen dreien möchte ich für die vielen wissenschaftlichen Diskussionen und Gespräche über das Leben an sich danken. Ohne euch hätte ich es so nicht geschafft.

Ein weiteres großes Dankeschön geht an Nadine für ihre exzellente Unterstützung im Labor, ihre wertvollen Verbesserungsvorschläge und ihre nicht zu erschütternde Fröhlichkeit. Außerdem möchte ich mich bei Axel und Martin bedanken, für die vielen spannenden Diskussionen über kleinere und größere Probleme im und ums Labor, mögen eure Turbopumpen lange halten. Georg und Nicola möchte ich vor allem dafür danken, dass sie die Freude am Arbeiten mit großen (und teuren) Maschinen in mir geweckt haben und mir die Scheu davor genommen haben, diesen nur mit einem Schraubenzieher „bewaffnet" gegenüberzutreten.

Auch die anderen Metabolomics-Freaks können hier natürlich nicht unerwähnt bleiben: Dr. Michael Gruber hat die einzigartige Gabe, alle Leute in seinem Umkreis zum Lächeln zu bringen, dafür ein herzliches „Mahlzeit!". Und schließlich unsere Projektstudenten/Diplomanden Stephan, Steffi und Magdalena, durch deren Fragen mir immer wieder bewusst wurde, wie viel es für mich weiterhin zu lernen gibt.

Auch bei allen anderen Kollegen am Institut für Funktionelle Genomik und dem KFB möchte ich mich sehr herzlich bedanken. Stellvertretend seien hier genannt: Sabine, das Organisationsgenie; Marian, für sein rasches und exzellentes Lösen aller Computer-bezogenen Probleme; Thomas, als treuester Mensakollege; den Organisatoren der großartigen Party; und auch allen anderen.

Mein größter Dank geht an meine Familie, für ihre Unterstützung in allen meinen Entscheidungen und Lebenslagen und für die Gewissheit, dass es für mich immer einen Platz geben wird, an den ich zurückkehren kann und an Stephan, seine Unterstützung und Liebe haben mein Leben aufs Großartigste bereichert.

2 Abbreviations and Acronyms

2-KG	2-ketoglutarate
6PG	6-phosphoglycerate
Aco	cis-aconitate
ADC	analog-to-digital converter
ADP	adenosine diphosphate
AEC	anion-exchange chromatography
AMP	adenosine monophosphate
BGE	background electrolyte
cdw	cell dry weight
CE	capillary electrophoresis
Cit	citrate
CoA	coenzyme A
CTL	cytotoxic T-lymphocytes
CZE	capillary zone electrophoresis
DHAP	dihydroxyacetone phosphate
ED	entner doudoroff
EI	electron impact ionization
EIC	extracted ion chromatogram
EOF	electroosmotic flow
ESI	electrospray ionization
F6P	fructose-6-phosphate
FBP	fructose-1,6-bisphosphate
FTICR	fourier-transform ion cyclotron resonance
Fum	fumarate
FWHM	full width at half maximum
G1P	glucose-1-phosphate
G3P	glyceraldehydes-3-phosphate
G6P	glucose-6-phosphate
GC	gas chromatography

glycerol-1P	glycerol-1-phosphate
Glyox	glyoxylate
GSH	glutathione
H	height of a theoretical plate
HILIC	hydrophilic interaction liquid chromatography
HPLC	high-performance liquid chromatography
IP	ion pair
IS	internal standard
Isocit	isocitrate
IT	ion trap
Lac	lactate
LB	Luria-Bertani
LC	liquid chromatography
LG	lactoylglutathione
LIT	linear ion trap
LLOQ	lower limit of quantification
LOD	limit of detection
LOQ	limit of quantification
M9	minimal medium
Mal	malate
MG	methylglyoxal
MRM	multiple reaction monitoring
MS	mass spectrometry / mass spectrometer
MS/MS	tandem mass spectrometry
MT	migration time
NAD	nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
NMR	nuclear magnetic resonance
NP	normal phase
OAA	oxaloacetate
PARAFAC	parallel factor analysis

PCA	principal component analysis
PEP	phosphoenolpyruvate
PESU	polyethersulfonate
PG	phosphoglycerate
PGC	porous graphitic carbon
PGP	1,3-bisphosphoglycerate
PIPES	1,4-piperazinediethanesulfonic acid
PPP	pentose phosphate pathway
Pyr	pyruvate
QC	quality control
QTRAP	triple quadrupole – linear ion trap hybrid mass spectrometer
r	correlation coefficient
R5P	ribose-5-phosphate
RF	radio frequency
RP	reversed phase
RR	rapid resolution
RSD	relative standard deviation
RSQ	square of the correlation coefficient r
RT	retention time
Ru5P	ribulose-5-phosphate
S7P	sedoheptulose-7-phosphate
SIM	single ion monitoring
SQ	single quadrupole
SRM	single reaction monitoring
Succ	succinate
TBA	tributylamine
TCA	tricarboxylic acid cycle / citric acid cycle
TDC	time-to-digital converter
TIC	total ion chromatogram
TOF	time-of-flight
TQ	triple quadrupole

ULOQ	upper limit of quantification
UPLC	ultra-performance liquid chromatography
wt	wild type
X5P	xylulose-5-phosphate

3 Motivation

The 'metabolome'[1] represents the nonstructural molecular phenotype of a cell, tissue or organism. It comprises a great number of compounds of low molecular mass (typically <1,000 Da), which differ greatly in their chemical and physical properties.[2] Metabolomics denotes the comprehensive and quantitative analysis of all the small molecules present in the system. However, to date no single method has been described that is suitable for an exhaustive metabolomic analysis.[2] One strategy to circumvent this problem is the integration of different methods into a comprehensive analysis platform. Most methods for metabolome analysis either aim at the accurate quantification of a selected subset of metabolites or at the general profiling of a great number of compounds with very limited information on abundances. Another important prerequisite in metabolomics is the capability to measure a high number of samples in a reasonable amount of time, in order to account for biological and analytical variability. For this purpose, it is not practicable to analyze one sample with many different methods. The integration of these general conditions points to a substantial need for methods that integrate several analytical aspects. Examples are the combined analysis of different metabolite classes (e.g. amino acids and fatty acids), the combination of accurate quantification with screening (see Aim # 1) or the integration of highly sensitive quantification with structural information (see Aim # 3).

Aim # 1: Development of a method that combines accurate quantification of metabolites of the central carbon metabolism with screening for unknowns

One major goal of the present work was the development of a method that would allow both the absolute quantitation of selected metabolites and the semiquantitative screening for unknown analytes and their identification with high confidence. The initial focus was on the analysis of the central carbon metabolism, an important node for the energy metabolism of a cell that

generates many precursors for diverse cellular functions, such as amino acids and nucleotides (Figure 1). The metabolites involved in the corresponding pathways are mostly negatively charged, and many are phosphorylated. These small, hydrophilic molecules are not well suited for reversed-phase liquid chromatography – mass spectrometry (RP-LC-MS) analysis.[3-5] On the other hand, gas chromatography (GC), which is the method of choice for many global metabolome studies,[6-9] requires pre-column derivatization and is not applicable to the quantitation of thermally instable compounds, such as phosphorylated metabolites.

Therefore, capillary electrophoresis (CE) was chosen for the separation of the charged metabolites, because it yields narrow peaks and has a high separation efficiency for isobaric compounds such as hexose-phosphates.[10] A time-of-flight mass spectrometer (TOF-MS) was used for analysis, because it provides a wide dynamic range for absolute quantification and high mass accuracy and fast scan speed over a wide mass range, which are prerequisites for screening applications. The CE-TOF-MS method was optimized with regard to the separation of metabolically important isobars, e.g. hexose-phosphates and glycerol-phosphates, method stability and robust quantification. Method parameters such as limit of detection (LOD), limits of quantification (LOQ), linear range and reproducibilities were established.
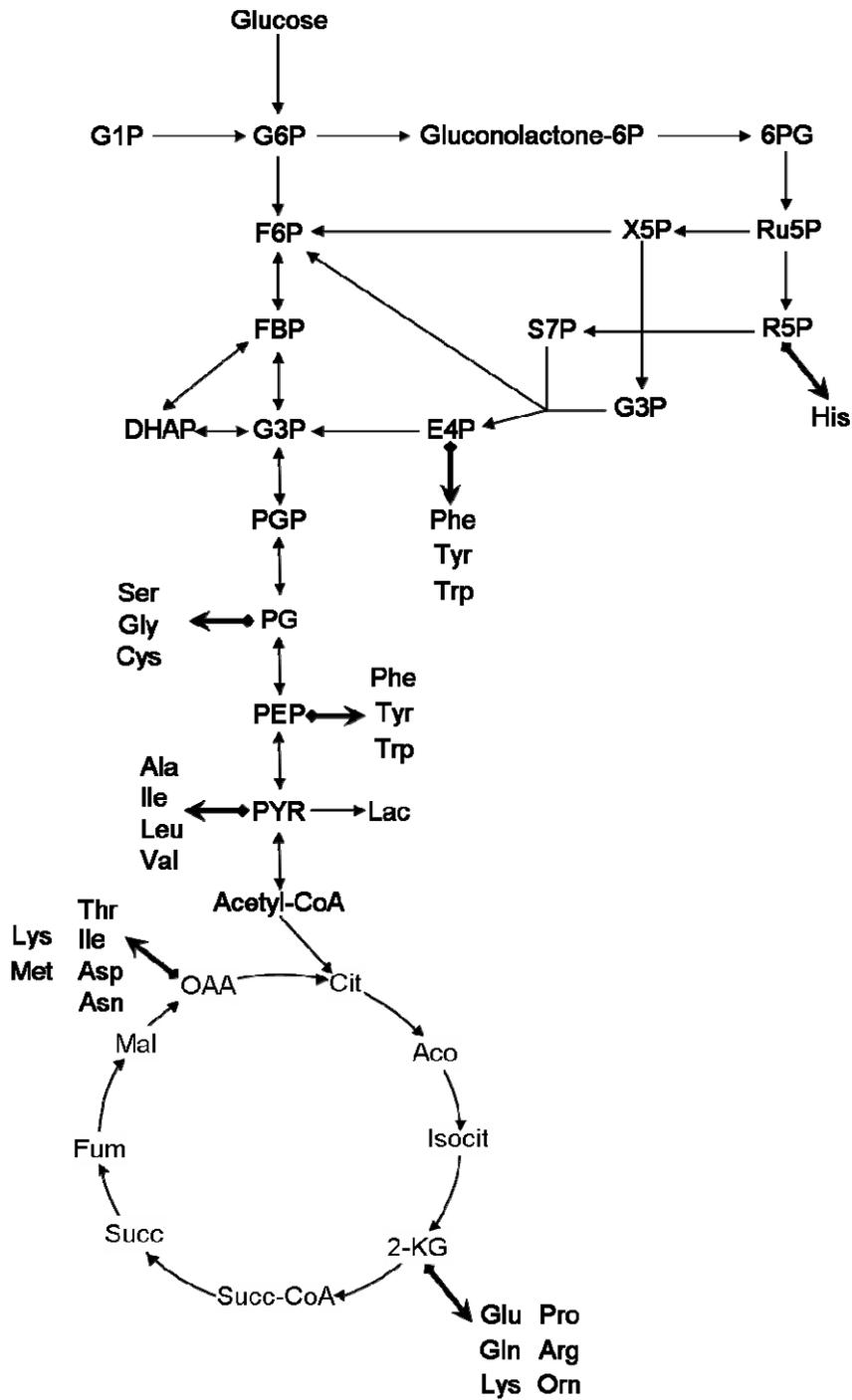
Figure 1: Schematics of the central carbon metabolism. Arrows denote the respective amino acids that are built from the indicated precursors.

Aim # 2: Application of CE-TOF-MS to the metabolome analysis of *E. coli*

The CE-TOF-MS method was applied to the metabolomic analysis of the model organism *Escherichia coli*. To that end, several sample preparation steps had to be optimized, including cell harvesting, extraction of metabolites, and the evaluation of matrix effects during CE-MS analysis. The method was then applied to the analysis of two different *E. coli* strains that differ in glycolytic flux distribution. We quantitated a selected subset of central carbon metabolites and identified many more compounds. We adapted programs for peak detection and alignment to our specific needs and compared their performance with data from the absolute quantification. Furthermore, we introduced several measures to ascertain the identity of every analyte. Finally, the semi-quantitative information from the high-confidence identification of additional metabolites was integrated with the data on absolute levels of selected metabolites and, thus, we were able to obtain a comprehensive picture of metabolic differences between the two *E. coli* strains studied.

Aim # 3: Development of a method that combines sensitive quantification with structural information

The CE-TOF-MS method was well suited for the application described above. However, it lacked robustness for the study of large sample batches and its detection limits were mostly in the low micromolar range. These considerations prompted us to develop a method that allowed for the robust quantification of target analytes with low limits of detection, yet also yielded information on molecular structure. We chose ion-pair liquid chromatography, which had been shown to yield good separation of negatively charged metabolites.[11] The LC column was directly coupled to a triple quadrupole ion trap mass spectrometer operated in single reaction monitoring (SRM) mode.

In order to reduce analysis time, a rapid resolution HPLC system on sub-2µm octadecyl silica partiles was implemented. Initially, the method was tailored for the rapid quantitative measurement of lactate and pyruvate. Later the method was expanded to a panel of metabolites representing both the upper (glycoly-

sis, PPP, methylglyoxal pathway) and the lower part (TCA, glyoxylate shunt) of the central carbon metabolism.

Aim # 4: A simple and rapid method for flux analysis in *E. coli* and human cells

Measuring metabolite concentrations yields interesting information on biological systems. However, in order to access knowledge on perturbation in intracellular fluxes, the fate of labeled precursors has to be monitored. The applicability of the optimized RR-IP-LC-MS/MS method for flux analysis was therefore evaluated. Alterations in fluxes of the central carbon metabolism were monitored by the conversion of $[1,2-^{13}C_2]$glucose to lactate and other intermediates of important metabolic pathways. Again *E. coli* was used as a model organism. In an initial proof-of-principle study, we measured changes in the ratio of glycolysis to pentose phosphate pathway (PPP) in a wild type vs. a mutant *E. coli* strain to confirm that our method would match results obtained previously by gas chromatography-mass spectrometry.[12] In a next step we evaluated the suitability of this strategy for the analysis of other fluxes, such as the citric acid cycle (TCA) and the glyoxylate shunt. Finally, the method was extended to human immune cells to elucidate the effect of high extracellular amounts of lactate, which are found in certain tumors, on immune cell function.

4 Background

An abbreviated version of this chapter was published in Analytical Bioanalytical Chemistry.[13]

4.1 Metabolomics

The complete set of small molecules synthesized by an organism, termed metabolome, [1] closest reflects the actual nonstructural phenotype and, thus, the effective behavior of the cellular system of interest (Figure 2).
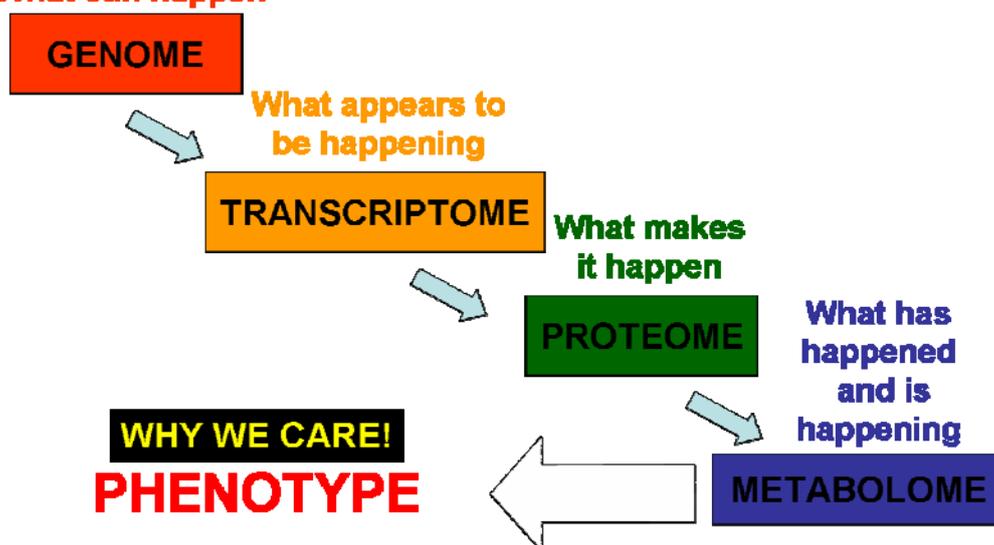
Figure 2: The 'Omics' cascade – Information flow in a cell. Adapted from [2]

The metabolite levels are constantly influenced by enzymatic reactions catalyzed by proteins, which themselves are highly regulated by transcriptional and translational control and by interactions with other proteins and metabolites. One can therefore expect that the measurement of global metabolite levels will reflect the actual state of a certain biological system under defined circumstances. However, the term 'metabolome' comprises a great number of compounds (e.g. an estimated number of > 200,000 in the plant kingdom [14]) that share the property of a low molecular mass (usually < 1,000 Da) but differ

in many other chemical and physical properties. So far, there exists no single method to identify and quantify all of them.[2]

Different strategies have emerged for tackling the complex data space of the metabolome. However, they differ in their demand on prior knowledge about the analytes of interest, and also require methods with different capabilities for the identification and quantification of the analytes.[2, 15, 16]

Metabolic profiling focuses on groups of metabolites that are selected by their relation to a specific metabolic pathway, their compound class or some other shared property. This allows the analysis to be tailored to the specific features of these metabolites. Usually, the focus during method development is on robust quantification and specificity. This approach is very useful for hypothesis-testing and adressing well-defined biological questions.

Target analysis is even more directed than metabolic profiling and only very few analytes are measured. They are often directly related to a genetic perturbation, such as substrates or products of enzymatic reactions, or they serve as biomarkers for a certain disease.[17]

However, these approaches are not truly 'omic' in the sense of yielding a global picture of the metabolic state of the system of interest. Prior knowledge about the nature of the analytes is required and all information on metabolites not specifically included in the analysis is lost.

Metabolic fingerprinting, on the contrary, aims at the detection of as many analytes as possible. Initially, the focus of this approach was mainly on the detection of patterns, the so called "fingerprints", that were caused by perturbations in the metabolism, such as mutations, disease or exposure to chemotherapeutics or toxins (Figure 3).
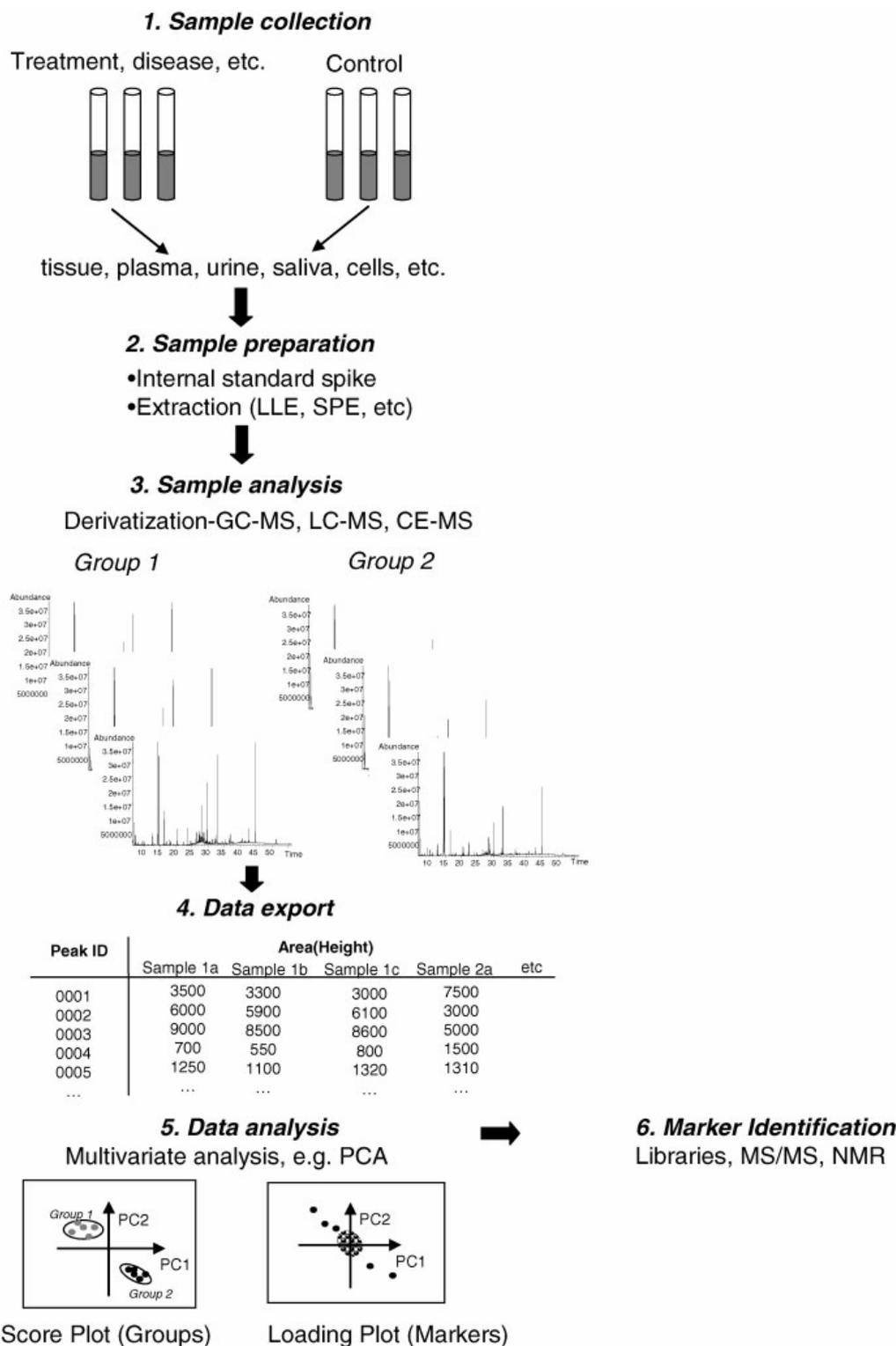
Figure 3: Simplified workflow for a metabolic fingerprinting analysis. Taken from [2]

The reasoning is, that those patterns can later be used as markers for these perturbations, without immediate need for the identification of its constituents. However, information on the metabolites that are altered under certain conditions is likely to elucidate the mechanism underlying the perturbation. Therefore, it will be necessary to develop screening methods that yield a high amount of additional information on the detected compounds, such as accurate mass, isotopic pattern, structural information and other physical or chemical properties.

A drawback of this approach is the relative quantification of peaks between different samples, i.e. no absolute concentrations are determined. It is thus very important to implement quality control steps and randomization to prevent wrong conclusions caused by changes in instrument and sample stability.

Metabolic footprinting uses the same methods as fingerprinting but is limited to the analysis of metabolites in cell culture media. The reasoning is, that compounds excreted by a cell or taken up from the medium will also give valuable insights into a cell's phenotype and physiological state.[18] A major advantage is, that samples can be analyzed directly without the need for quenching, i.e. the rapid stop of enzymatic acitivity, which often introduces additional variation or bias, as discussed in chapter 4.4.

4.2 Detection methods for metabolomics

The ultimate goal of metabolomics is the comprehensive analysis of all small molecules in a given system. Even though elaborate separation methods are usually employed, the use of a detector that can provide additional information on each detected compound is of high value. The detection methods mainly used for metabolomics are mass spectrometry and NMR. Mass spectrometry yields information on the amount of analyte present, its mass and its structure, depending on the type of mass analyzer used. NMR also produces quantitative data and is an excellent tool for structure elucidation.

4.2.1 Mass spectrometry

Mass spectrometry is commonly used in metabolomic studies due to its high sensitivity and throughput and the ability to confirm or even establish the identity of the detected compounds. Today, a number of different types of mass spectrometers are on the market and it is very important to understand their power and limitations in order to choose the best detector for the problem at hand.

Time-of-Flight (TOF)

This detector is based on the observation, that the time ions need to travel a fixed distance depends on their m/z ratio when accelerated with the same amount of energy (Figure 4).

The kinetic energy of ions leaving the source

$$T = eV = \frac{mv^2}{2}$$

The ion velocity equals the flight path (L) divided by time

$$v = \frac{L}{t}$$

The flight time is proportional to the square root of the mass, which makes the mass proportional to t^2

$$t = L\sqrt{\frac{m}{e}\frac{1}{2V}}$$

$$m/z = [2V/L^2]t^2$$
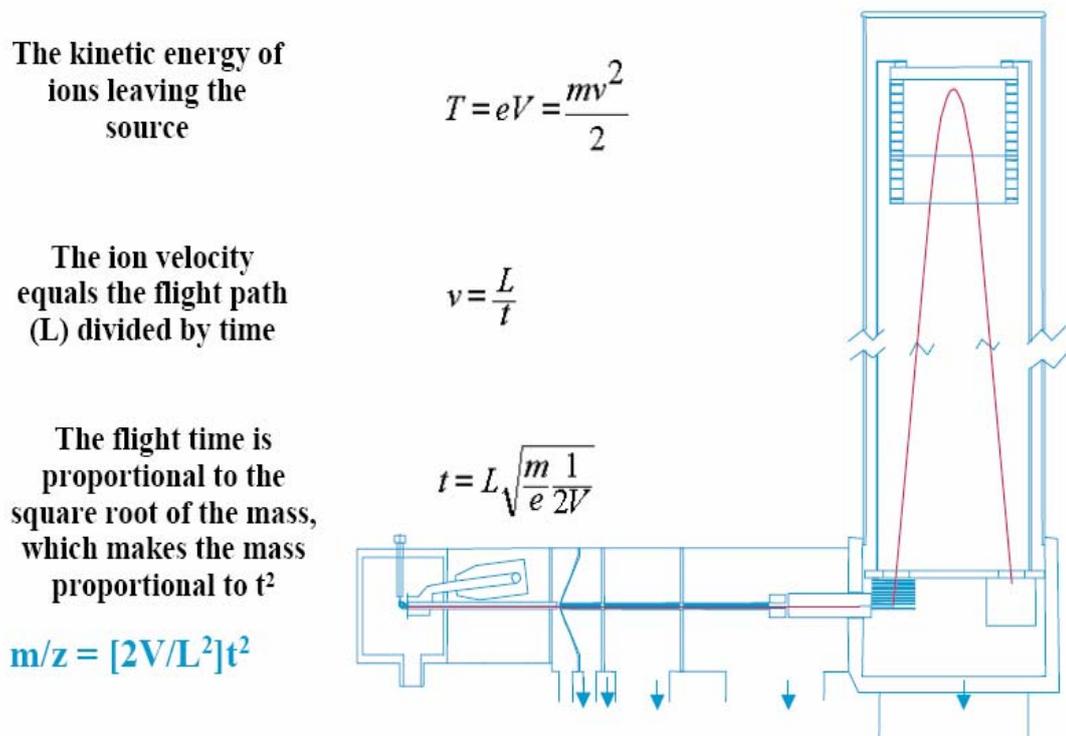
Figure 4: Schematics of a TOF-MS with reflectron (from [19])

Modern TOF instruments are equipped with a curved-field reflectron in the flight tube, which corrects for initial differences in kinetic energy.[20] The newest

generation of TOF instruments contains an analog-to-digital converter (ADC) that provides linearity over four orders of magnitude. This is due to the fact that the ADC records the total signal strength at a very high rate (1-4 GHz) as compared to the previously used time-to-digital converter (TDC), which measured the exact arrival time of the first ion of a given mass.[21]

The advantages of the TOF-MS are its high scan speed, making it the fastest MS analyzer available. Also, it has the highest practical mass range of all analyzers, which is, however, limited if a reflectron is used.[22] The new generation of TOF instruments achieves a very high mass accuracy (down to 3-5 ppm), which makes it very suitable for screening purposes, because identification is more specific.[23] The wide linear range allows the accurate quantification of isotope peaks. The combination of accurate mass and exact quantification of isotope distributions is used for theoretical molecular formula generation,[24] which is useful in the identification of unknown compounds. The main limitation of the TOF analyzer is its higher detection limits in comparison to a triple quadrupole instrument, because no precursor selection and, thus, background reduction can be performed. In addition, the instrument is sensitive to temperature changes and in order to achieve the highest mass accuracies, re-calibration with a standard mix is required regularly.

Quadrupole instruments

A quadrupole ion filter consists of four parallel rods. The rods have a fixed DC (direct current) and alternating RF (radio-frequency) voltages applied to them. Depending on the produced electric field, only ions of a particular m/z will be allowed to pass, all the other ions will be deflected into the rods.

The quadrupole ion filter is the most common mass analyzer and is used in routine analysis, for which it is well suited because of its good reproducibility and excellent stability.[22] One disadvantage of the quadrupole analyzer is its unit mass resolution. For that reason, it is not well suited for identification of unknowns. Its mass range is limited; it is able to analyze ions up to an m/z ratio of about 3,000-4,000. It has been used in metabolomics, albeit multiple

analyses had to be performed on each sample to cover the entire mass range with sufficient sensitivity.[25]

In triple quadrupole (TQ) instruments, three quadrupoles are combined sequentially. The second quadrupole (Q2) functions as a collision cell, where analytes selected in Q1 are fragmented. These fragments are selected in Q3 and analyzed. Several modes of tandem mass spectrometry (MS/MS) are available (Figure 5). The TQ is the instrument of choice if robust and sensitive quantification is required.[26] Also some information on molecular structure can be gained from product ion scans.
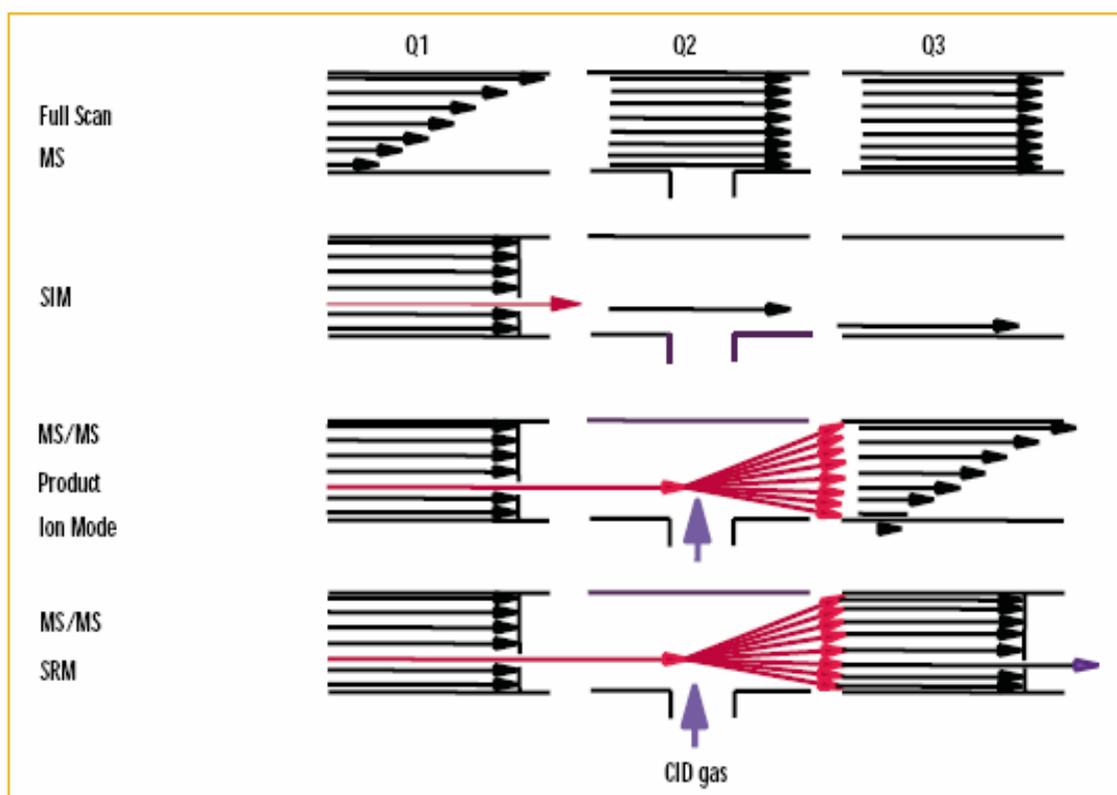
Figure 5: Scan modes in a TQ instrument (from [27]).

Quadrupolar or 3D ion trap

The ion trap (IT) consists of a ring electrode and two endcap electrodes. An AC potential of constant frequency and variable amplitude is applied to the ring electrode to produce a 3D quadrupolar potential field within the trapping cavity, which traps ions in a stable oscillating trajectory. During detection, the

electrode system potentials are altered to produce instabilities in the ion trajectories and thus eject the ions according to their m/z ratio in axial direction.

The advantages of the ion trap are its high sensitivity due to accumulation of ions within the trap and the provision of multiple sequential MS/MS experiments (MS[n]) that facilitate structure elucidation of larger molecules, *e.g.* peptide sequences. The limitations of the ion trap are poor quantitation, low mass accuracy and a poor dynamic range. The trapped ions are subject to space charge effects and ion molecule reactions, thereby complicating data analysis.[28]

Linear ion trap

The linear ion trap (LIT) has a design similar to a quadrupole mass analyzer, with the addition that a static electrical potential can be applied on end electrodes to confine the ions axially. Compared to 3D-ion traps, the LIT has higher injection efficiencies and higher ion storage capacities.[29] It is often combined with other mass analyzers in hybrid instruments, as discussed in the next paragraphs. The LIT is mostly used for applications demanding low detection limits and great robustness, and for MS/MS and MS[n] characterization.[30, 31] A dynamic range of up to five orders of magnitude with a detection limit of ~10zmol (~6000 molecules) has been reported.[32]

Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICR)

In this type of mass analyzer, ions are trapped in a stable orbit in an electromagnetic field. A resonant radio frequency (RF) signal is used to excite the orbiting ions and as a result the ions produce a detectable image current. By a short RF impuls all ions with different m/z ratios are excited simultaneously and the evolution of the resulting signal in time is recorded. Numerical transformation from the time to the frequency domain, the so-called "Fourier transformation", provides a conventional spectrum. Due to the general ease to precisely measure frequencies, very high mass resolutions are achieved.[33] High

mass accuracies aid in the identification of unknown compounds, however it was shown that for larger molecules, additional information on isotope patterns is very important for reliable identification.[23] In addition, ion excitation can be used for MS/MS and MS^n experiments or for selective ion ejection.[33]

The limitations of FT-ICR are for one its limited dynamic range. Additionally, the mass spectrum is complicated by space charge effects and ion molecule reactions during ion storage and artefacts such as harmonics and sidebands. The handling of the FT-ICR is rather complex and it is very expensive in acquisition and maintenance.

Triple quadrupole – linear ion trap (QTRAP)

The QTRAP is essentially a triple quadrupole instrument, where the third quadrupole can also be used as a LIT. Its advantage in comparison to the TQ is the high sensitivity in scan mode, due to ion accumulation by trapping. Also, MS^n can be performed when the trap function is activated. The QTRAP is often used in MRM mode for the robust quantification of compounds, where very low limits of detection can be achieved.[11] In a successive experiment, the same sample can be analyzed in trap mode and thus screened for additional analytes of interest. In addition, high-sensitivity product ion scans can be used as an identification tool by investigating characteristic fragmentation patterns.

Quadrupole time-of-flight (QTOF)

The QTOF is a combination of two quadrupoles with a TOF mass analyzer, which will scan all the ions emerging from the collision cell. The Q1 can be used as a mass filter to substantially reduce background noise and, thus, achieve higher sensitivity. On the other hand, the TOF can be used to acquire a full product ion spectrum of the selected and fragmented ion or it can be used for screening. The advantage of this hybrid instrument type over a normal TOF is its ability to provide accurate mass measurements together with structural information.[26, 34] As a consequence, however, detection limits are not as low as in TQ-MS.[35]

4.2.2 Nuclear magnetic resonance spectroscopy (NMR)

Atomic nuclei with non-zero spin adopt quantized orientations in magnetic fields. Transitions between these quantized states can be induced by radiofrequency irradiation. The transition energies are dependent on the type of the atom and they are additionally modulated by the electronic environment of each respective nucleus, which is the origin of the different chemical shift values of atoms occurring in different molecular positions.[36] Analysis by NMR is based on (i) the modulation of the resonance energy by the electronic enivonment of a given atomic nucleus and (ii) the mutual influence that neighbouring nuclei exert on each other via chemical bonds (scalar coupling) and/or by interaction through space (nuclear Overhauser effect, NOE).[36]

The advantages of NMR are its capability for the structural characterization of unknown compounds and its robust quantification, at least in 1D-NMR.[37] However, its requirement for large sample volumes adversely affects sensitivity. LODs are under optimal conditions in the high nanomolar to low micromolar range for small molecules. Yet they depend very much on the spectral background, the solvent and the sample matrix.[38] The linear dynamic range of conventional instruments is roughly three orders of magnitude. Moreover, certain method parameters such as intersample peak position and line width are very sensitive to the chemical environment of the sample (e.g. pH, ion strength, etc.) and external factors such as tempterature.[39]

There are numerous applications for NMR in metabolic analyses. It is often used for structure elucidation of complex secondary plant metabolites.[36] NMR is also used for pathway elucidation, where the fate of isotope tracers is monitored.[36] Most metabolomics studies to date use NMR for metabolite profiling without absolute quantification. NMR has been used to investigate the effects of gender, diurnal variation and age in human urinary samples. Accuracy and precision from the measurement by three analysts was evaluated.[37] Several studies using NMR for nutritional applications are reviewed in.[39]

Figure 6: (a) One-dimensional ^{1}H NMR spectrum of an equimolar mixture of the 26 small-molecule standards listed in Supporting Information Table 1. (b) Two-dimensional ^{1}H-^{13}C HSQC NMR spectra of the same synthetic mixture (red) overlaid onto a spectrum of aqueous whole-plant extract from *A. thaliana* (blue). Provided by [38].

Quantification of complex biological samples by NMR is still in its infancy. The problem with analyzing complex spectra is the fact that overlapping peaks do not scale in the discrete linear fashion that typifies well-isolated peaks.[38] This impediment can be circumvented by applying two-dimensional (2D) ^{1}H-^{13}C NMR strategies (Figure 6). However, in this approach cross-peak intensities are influenced by a greater number of variables, such as uneven excitation,

16

nonuniform relaxation, varying J-coupling values etc. Nevertheless, it was shown that 2D-NMR leads to superior quantification of complex samples over a concentration range of 1 – 30 mM. Metabolites from different organisms, e.g. *A. thaliana* and *S. cerevisiae*, were quantified by this approach.[38] The quantitative capabilities of a "targeted profiling" approach were evaluated by measuring simplified synthetic urine. It was found that this method is highly stable in PCA-based pattern recognition, insensitive to water suppression, relaxation times and scaling factors. Also, it was validated against various metabolites at physiological concentrations (9 µM – 8 mM).[40]

4.3 Methods for metabolome analysis

4.3.1 Direct infusion

Direct injection of samples into mass spectrometers without any prior separation of analytes provides rapid, high-throughput analysis of crude samples or sample extracts. It is most often used with ESI-MS analysis, where the sample is delivered to the mass spectrometer at low µL/min flow rates. Usually, a high-resolution mass detector, such as a TOF-MS or a FTICR-MS is used,[2] because the only molecular feature for analyte identification is its mass. Thus, the higher the mass accuracy for the measurement of the metabolite of interest, the more reliable its identification will be. An alternative is using fragmentation patterns derived from MS/MS analysis. Obvious limitations of this approach are that it is (i) not able to discriminate between isobaric metabolites and (ii) matrix effects are likely to impede quantification and identification, e.g. by adduct formation and ion suppression.[41] Nevertheless, its suitability for very high-throughput analyses makes it an interesting tool in medical screening. To date the most important application of this approach is the screening for inborn errors of metabolism, especially for newborn screening. Also, the rapid profiling of secondary metabolites in plants and fungi has been described.[15]

4.3.2 Matrix-assisted laser desoption/ionization – mass spectrometry (MALDI-MS)

MALDI-MS is commonly used in proteomics. It has not been used until recently for the analysis of small molecules. This is mainly due to the strong interference from matrix ions in the low molecular weight range (< 500 m/z). In addition, the conventional matrices are much more suitable for ionization in positive mode, which is not feasible for small organic acids and phosphorylated compounds, which constitute some of the most important metabolites. The introduction of 9-aminoacridine (9-AA) as a matrix has been shown to yield low background and high sensitivity in negative mode.[42, 43] 9-AA was used for the analysis of metabolites in *E. coli* and islets of Langerhans by MALDI-TOF-MS.[44] LODs were between 15 nM and 1 µM. In *E. coli*, 60 metabolites were detected, 39 of which matched compounds in the MetaCyc database. Reproducibilities (RSD) for repeated measurements of one sample were acceptable with 8-12%. However, inter- sample reproducibility was worse (46-73%). A linear response for spiking one analyte into a metabolite mixture was seen over one order of magnitude. However, linearity was not good (0.94-0.988) and reproducibility of replicate measurements was rather poor (often > 100%).[45] To assess the quantitative capabilities of MALDI-MS in more detail, a metabolite cocktail was spiked with different concentrations of single metabolites.[46] The variations for replicate measurements were again rather high, which the authors contributed to the heterogeneity of analyte distribution on the surface. Nevertheless, it was clearly seen that changes in the concentration of one analyte often influenced the quantification of other metabolites. A correlation between chemical similarity and degree of influence was clearly established. These findings demonstrated the limited utility of MALDI-MS for quantitative analysis.

4.3.3 Capillary electrophoresis

In capillary zone electrophoresis (CZE) analytes are separated according to their mobility in an electric field. The mobility is influenced by the ion radius,

the charge of the molecule and the viscosity of the solution (1). The electrophoretic mobility is denoted as μ_p, the charge of the analyte as z, viscosity of the solution as η and r stands for the ion radius.

$$\mu_p = \frac{z}{6\pi\eta r}$$

(1)

The effective mobility in a given system is in addition influenced by the electroosmotic flow (EOF). The EOF depends on the chemical nature of the capillary wall, as well as on pH and ion strength of the background electrolyte (BGE). In capillary electrophoresis, metabolites can be measured directly in aqueous solution, eliminating the need for time-consuming sample preparation and derivatization steps. Peak capacities in CE are very high, due to the narrow peaks caused by the plug flow of the bulk liquid as compared to the parabolic flow profile in LC. Another advantage is that only minute amounts of sample, usually a few nanoliters, are required for one injection, which is important for samples of limited availability, e.g. patient derived samples, and permits the re-analysis of samples. However, the small injection volumes can also pose a disadvantage, in that they increase detection limits with regard to detectable concentrations. Another issue in CE is sample ion strength. While a low ion strength can increase resolution by stacking effect, [47, 48] high ion strength can lead to peak broadening and alterations in migration behavior.

A particular challenge in the analysis of intermediates of the central carbon metabolism is the presence of isobars such as glucose-6-phosphate (G6P) and fructose-6-phosphate (F6P) that also have similar mass spectrometric fragmentation patterns. These compounds require electrophoretic or chromatographic separation prior to mass spectrometric detection. Capillary electrophoresis has been successfully employed in the separation of isobaric compounds using coated capillaries and a reversed EOF.[25] By coupling CE to a quadrupole mass analyzer, methods for the analysis of positively and negatively charged central carbon metabolites and associated nucleotides were developed. With this method combination, more than 1,600 potential metabo-

lites were detected in *B. subtilis*.[10] However, the main disadvantage of this approach was the need for multiple analyses of each sample, because narrow m/z windows had to be analyzed in order to screen for unknown metabolites with high sensitivities.

Therefore, the coupling of CE to a time-of-flight (TOF) MS was introduced.[49] This allows for the scanning of the complete mass range of small molecules, e.g. 50 – 1000 m/z, several times per second and, thus, the detection of a great number of features that represent potential metabolites. The method was applied to the analysis of oxidative stress in liver cells and 1,859 peaks could be detected in mouse liver extracts. Comparative quantitative analysis between normal cells and liver cells that were oxidatively stressed by acetaminophen revealed multiple changes in metabolite levels, including the activation of the ophthalmate biosynthesis pathway.

The application of CE coupled to a triple quadrupole mass analyzer has not been demonstrated yet for the study of the central carbon metabolism, but its general utility has been demonstrated for the quantitative analysis of amino acids in human urine.[50] Highly selective multiple-reaction monitoring (MRM) in positive ion mode was used to quantify 32 free amino acids. LODs ranged between 0.1 and 14 µM, the calibration was linear between 10 – 200 µM. Apparently, the relatively high LODs are mainly due to the small volumes injected in CE.

4.3.4 Liquid chromatography

Reverse-phase (RP) LC is usually the method of choice for direct coupling to a mass spectrometer. It is robust, yields stable retention times and uses solvents that are highly compatible with ESI-MS. However, many important metabolites are highly polar or even charged and, therefore, do not interact with the hydrophobic alkyl chains of a RP surface. Different approaches were suggested to overcome this problem, with the most promising being ion-pair liquid chromatography (IP-LC) and hydrophilic interaction chromatography (HILIC).

4.3.4.1 Reversed-phase liquid chromatography

Recently, RP-LC has been used in metabolomics following the introduction of modifications that served to increase the otherwise poor resolution of polar metabolites. Analysis of polar analytes in RP-mode often requires starting the LC-gradient with 100% water to achieve analyte retention. However, the brush-like structure of conventional C18 phases will collapse under these conditions and analyte retention is impeded. In addition, pore-dewetting occurs, which denotes the exclusion of the highly polar solvent from narrow RP-coated pores. Moderately polar compounds can be separated on polar- or hydrophilic-endcapped C18 phases. The endcapping groups allow the silica surface to be wetted with water and thus enable the full interaction with the longer alkyl chains. Another method to use C18 columns with highly aqueous mobile phases is using polar-embedded stationary phases. Due to the incorporation of a polar functional group in the alkyl ligand close to the surface of the silica gel, the phase remains solvated even at 100% water.[51]

One strategy to use RP-LC for small, highly polar molecules is the employment of alternative stationary phases or support material, such as monolithic phases and porous graphitic carbon columns. The advantage of monolithic silica capillary columns is their up to 30 times greater permeability as compared to 5-µm particle columns, due to the much higher porosity of the silica rod.[52] A monolithic RP column of 90 cm length was applied for the analysis of the *Arabidopsis thaliana* metabolome using an ion trap MS in both positive and negative mode with continuous polarity switching. Several hundred peaks were observed and the high resolution obtained with the long column led to a reduction of noise and thus better quantitation. Monolithic columns are often used for high-speed separations.[53] Especially effective is the direct coupling to high-flow on-line extraction. Cycle times, i.e. extraction, separation and analysis, of less than two minutes were achieved by these approaches.[54] Another interesting development is the use of monolithic columns as the second dimension in online 2-D LC-MS separations. The high speed that can be

achieved on these columns allows the direct analysis of fractions collected from the high-resolution separation employed in the first dimension.[55]

Porous graphitic carbon (PGC) stationary phases show high retention of analytes with high polarity. The retention mechanism is determined by hydrophobic eluent-analyte interactions and electronic interactions of polarizable functional groups in the analyte with the delocalized π-electrons of the graphite surface.[56] They are used for the separation of phosphorylated carbohydrates, where high resolution of isomers can be achieved.[57] However, they did not prove suitable for the separation of nucleotides, which had to be analyzed with a different method. Nucleotides and nucleosides were later separated on a PGC column after extensive optimization of the LC parameters.[58] The use of acetonitrile instead of methanol and ammonium acetate and diethylamine as modifiers finaly yielded optimal retention and peak shapes. A method for the analysis of sugars and sugar phosphates was developed on a PGC column coupled to an ion trap MS that achieved good linearity between 1-100 µM with LODs of 0.1-1.5 µM.[59] However, a complex gradient with three solvents had to be used to separate all analytes in a single run.

4.3.4.2 Rapid resolution – ultra performance liquid chromatography

An important goal in metabolomics – and other 'omics' approaches – is the analysis of as many analytes as possible in a short time. Thus, methods are required that yield high peak capacities, i.e. narrow peaks and high resolution. From basic chromatographic theory [60] it becomes clear that reducing particle size is one means to increase separation efficiency (Figure 7).

$$H = 2d_p + 6/u + d_p^2 u/20$$
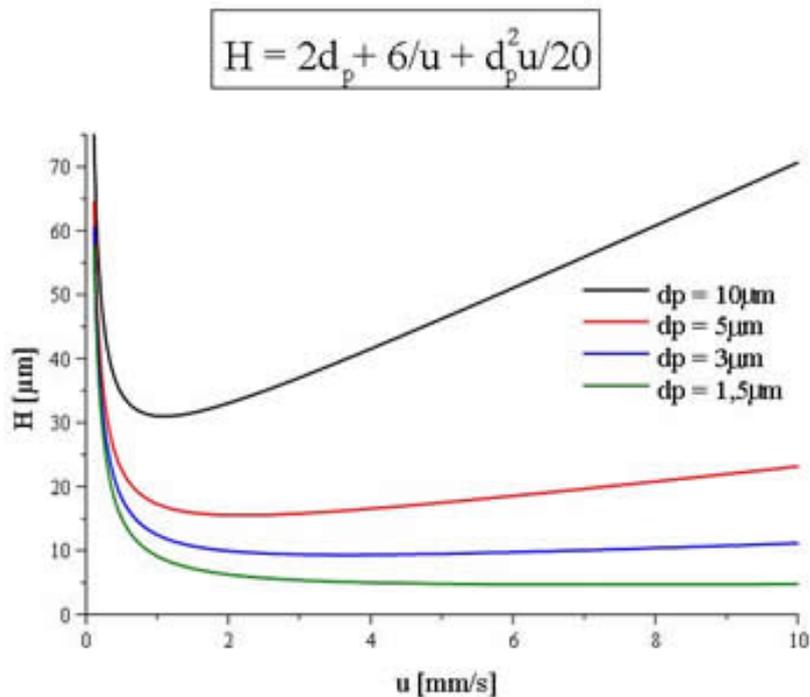
Figure 7: Van-Deemter Plot for different stationary phase particle sizes. Adapted from Halasz et al., 1975.

The Van Deemter curve for small particles indicates a low influence of the C-term on H when the flow rate is increased. This enables faster separations without loss of efficiency. Reducing the particle size by a factor of 2 doubles separation efficiency and, thus, increases the resolution by a factor of 1.4.[61] However, back-pressure is increased by a factor of 8. The introduction of LC systems that can generate and tolerate high pressure (600-1000 bar) and columns packed with sub-2µm particles led to the rise of commercialized approaches termed ultra-performance liquid chromatography (UPLC, Waters), and rapid resolution (RR, Agilent).

To date there exist a number of proof-of-principle studies demonstrating the higher peak capacities, increased resolution and/or shorter analysis times of UPLC as compared to conventional RP-LC.[62, 63] For example, the use of UPLC led to an increase in peak capacity by more than 2-fold, an almost 10-fold increase in speed and 3- to 5-fold lower detection limits as compared to a conventional 3.5-µm stationary phase. The number of peaks detected in a

typical mouse urine sample was increased from ~2,000 for HPLC-MS to ~10,000 for UPLC-MS. These features were subsequently used to distinguish different mouse populations by principal component analysis (PCA).[62]

A thorough comparison of 5, 3.5 and 1.7 μm particles under different conditions was performed by deVilliers et al. [61] Higher optimal velocities (~0.37 cm/s) and lower minimum plate heights (~4.4 μm) were found for well-retained compounds using 1.7-μm particles, as expected from theory. Experimental results demonstrated that the combination of high optimal flow rates and shorter column lengths allowed a gain in speed at constant efficiency by a factor of about 4.3 and 3.5 in comparison to 5 and 3.5 μm particles, respectively.

4.3.4.3 Ion-pair liquid chromatography

Many of the small molecules that comprise the metabolome are highly polar and charged. These compounds are not well suited for reversed-phase LC, which is, nevertheless, the method of choice for LC-MS coupling. One way of rendering charged analytes amenable to separation under RP conditions is the use of ion pair (IP) reagents. Different theories were proposed to describe the mechanism of IP-LC.[64] Charged hydrophobic species are used as IP-reagents. These compounds are adsorbed at the interface between the stationary and mobile phase, creating a charged surface, while the inorganic counterions form a corresponding diffuse layer. This implies that the IP-reagent creates an electrostatic surface potential, and that the magnitude of this potential is primarily determined by the surface concentration of the IP-reagent (Figure 8/B). Another hypothesis is the generation of uncharged complexes between the IP-reagent and the analyte, which are much less polar and will thus be retained on a C18 column, depicted in Figure 8/A.[64] The main parameters that influence the retention of an analyte in IP-LC are (i) the concentration of the IP-reagent in the solvent, (ii) the type of IP-reagent and (iii) the ionic strength of the mobile phase.
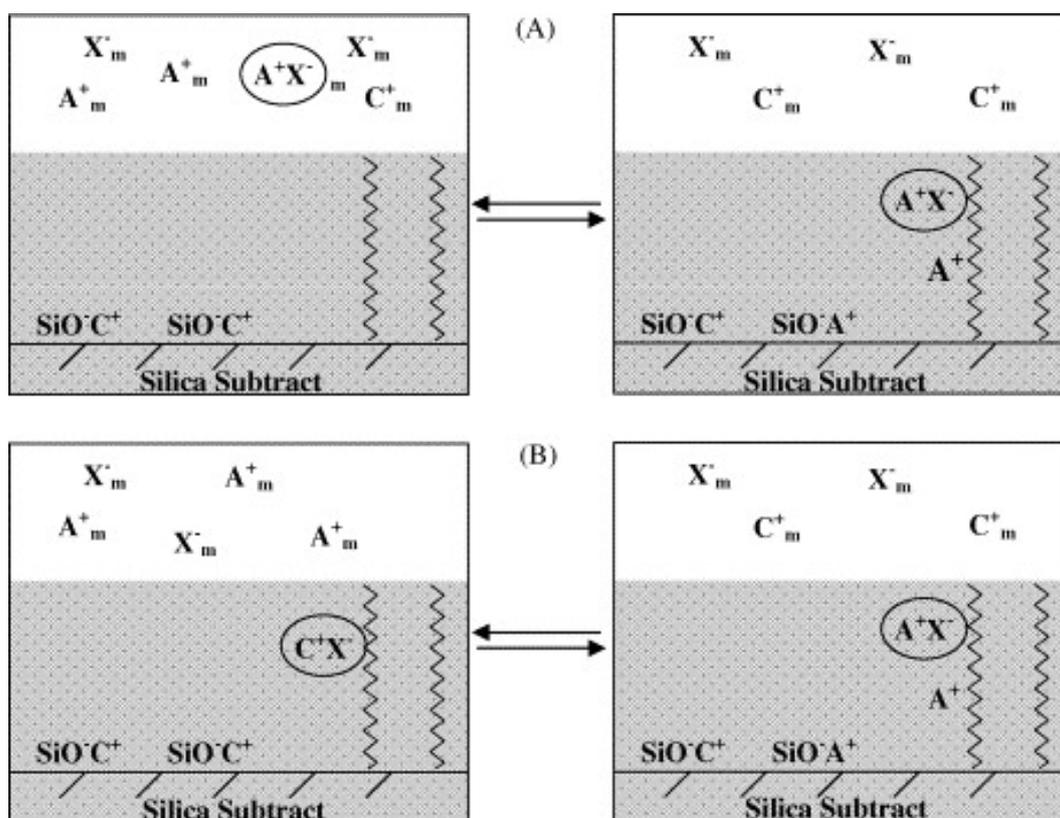
Figure 8: Cartoon illustrations of interactions between cationic solute A^+ and bonded silica phase according to: (A) "ion pair formation in the mobile phase" and (B) "dynamic ion-exchange in the stationary phase" mechanisms in the presence of a mobile phase additive X^-C^+. Taken from [65].

Different IP reagents have been used in the few studies on metabolome analysis reported to date. Hexylamine was used for the separation of nucleotides, coenzyme A and its esters as well as sugar nucleotides and sugar bisphosphates.[66] A combined methanol and pH gradient was employed for the separation. MS/MS analysis was performed on a 4000 QTRAP triple quadrupole ion trap MS. Imprecision was lower than 10% and good linearity was found over 3 orders of magnitude. LODs (S/N < 3) were between 0.1-1 ng injected on column. The method was applied to *E. coli*, *B. subtilis* and *L. plantarum*, with reproducibilities of multiple injections (n=12) below 10%. In another study, octylammonium acetate and an acetonitrile gradient were used to resolve intermediates of the pentose phosphate pathway (PPP). Phosphorylated sugars were measured in bloodspots, fibroblasts and lymphoblasts [67] on an API-3000 triple quadrupole MS operated in negative mode. Employing

multiple-reaction monitoring (MRM), intra-and inter-assay variation was between 3-17% and 5-21%, respectively, for the different biological matrices. Detection limits for most sugar phosphates were between 0.1-1 µM.

Recently, an IP-LC-MS/MS method was developed for the quantitative analysis of 29 negatively charged compounds of the central carbon metabolism, employing tributylammonium acetate (TBAA) as the volatile ion pair modifier.[11] The use of the 4000 QTRAP for monitoring selective MS/MS transitions for each compound resulted in very low LODs, mostly below 60 nM. Calibration curves were linear over three orders of magnitude, and relative standard deviations for repeated measurements of an *E. coli* cell extract were below 6% for all metabolites.[11]

4.3.4.4 Hydrophilic interaction liquid chromatography (HILIC)

In HILIC the separation of polar compounds is achieved by using a normal-phase (NP) stationary phase, with a solvent system similar to RP yet with a reversed gradient, i.e. from low to high water content. The weak eluent in HILIC is the organic solvent, mostly acetonitril, which contains a low percentage of water, typically 2.5%. The retention mechanism was suggested to involve mostly partitioning between the hydrophobic mobile phase and an aqueous layer formed on the stationary phase.[68] Therefore, contrary to RP-LC, unpolar analytes elute first from the HILIC column, while polar analytes are retained. The applicability of coupling HILIC to MS was demonstrated.[3] HILIC is also well suited for the separation of polar compounds that do not necessarily have to be charged under the separation conditions, such as sugars.[3] A comprehensive study on HILIC-MS of ~160 commercially available metabolites demonstrated its usefulness for diverse compound classes such as amino acids, amino acid derivatives, nucleosides, their bases and mono-, di- and tri-phosphates as well as CoA and its esters, carbohydrate derivatives and precursors, vitamins, carboxylic acids, redox-electron-carriers and others.[5] Of the 141 metabolites amenable to HILIC-MS/MS analysis, 69 could be quantified from *E. coli* extracts and for 39 significant changes were detected

upon carbon starvation. The method was subsequently used to investigate the metabolomic response to starvation in *E. coli* and *S. cerevisiae*.[69] The filter-culture approach used in this study allowed the very rapid quenching of the metabolism without risk of leakage (Figure 9). A very striking finding in this study was the dynamic range of the metabolite concentration changes upon starvation, with some metabolites decreasing or increasing by ~100 times. However, the median relative standard deviation of repeated analysis of the same extract was 13%, for duplicate independent experiments it was 31%.
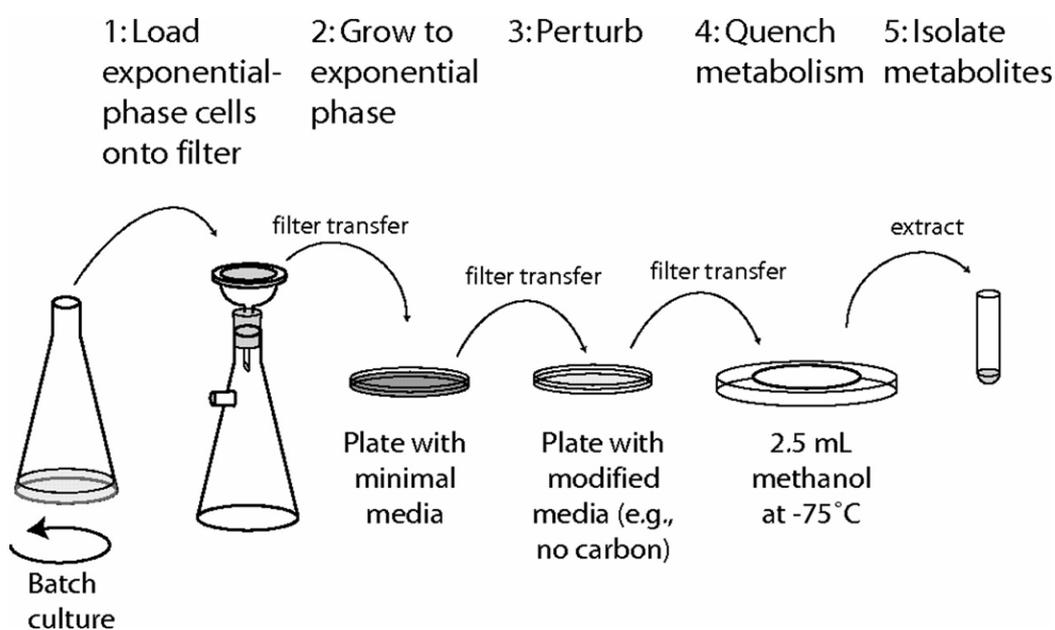
Figure 9: Schematic of the filter- culture approach. Taken from [69].

4.3.4.5 Anion-exchange liquid chromatography

Negatively charged analytes are amenable to anion-exchange chromatography (AEC), where they are usually eluted with increasing concentrations of sodium hydroxide and/or sodium acetate. However, the high salt concentrations are detrimental to mass spectrometric analysis. To solve this problem, Dionex Inc. introduced a postcolumn self-regenerating suppressor. This device electrochemically exchanges cations for protons.[70] AEC was used in several studies, mainly to investigate glycolytic intermediates, many of which were phosphorylated. AEC was coupled to a QQQ-MS in MRM mode with

negative ionization. LODs were in the range of 42-440 µM, with a linear range of about two orders of magnitude. Quantitative results were compared with an enzymatic assay and very good correlations were observed.[70] Several studies have described the combination of rapid sampling from a bioreactor for *S. cerevisae* and the subsequent analysis of glycolytic intermediates by AEC.[71-73] While the method seems to provide reasonable resolution for negatively charged metabolites, it suffers from poor sensitivity (LODs in the middle to high µM range) and limited dynamic range due to incomplete removal of ions in the suppressor.

4.3.5 Gas chromatography

4.3.5.1 GC-MS

Gas chromatography (GC) is still the method preferred for global metabolome studies.[6-8] It is very well suited for the analysis of polar small molecules upon derivatization, which is commonly done by silylation, provided the derivatives are stable at the high temperatures used in GC. GC-MS was used to quantify metabolites in different microorganisms.[74] Many different metabolite classes, such as alcohols, amines, fatty acids, organic acids, sugars and others, could be analyzed, with relative standard deviations mostly below 10%. LODs were 40-500 pg on-column, corresponding to 0.1-0.7 mmol/g dry weight. The method was applied to the analysis of *E. coli* samples harvested at different growth phases. Another application was the use of GC-MS for the evaluation of sample preparation methods in yeast.[75] Amino acids and organic acids were derivatized with methylchloroformate, sugars and related compounds as well as peptides were analyzed after oximation-silylation. The effect of quenching, extraction and water evaporation on metabolite recovery was assessed and the most suitable protocols were highlighted.

4.3.5.2 GC-TOF-MS

The introduction of GC-TOF-MS hyphenation allows the use of GC, a robust separation technique, for the screening and identification of unknown metabolites. In a proof-of-principle study, 77 metabolites found in *C. reinhardtii* under different conditions were relatively quantified.[7] Principal component analysis (PCA) was performed on 177 metabolites, both known and unknown, which allowed the clear separation of the different experimental groups. Later on, GC-TOF-MS was used for the discrimination of invasive ovarian carcinomas and ovarian borderline tumors based on their different metabolic patterns.[76] PCA and supervised predictive models allowed the separation of 88% of the borderline tumors from the carcinomas. However, only nine borderline tumors were analyzed in the study as compared to 66 invasive carcinomas.

4.3.5.3 GCxGC-TOF-MS

Comprehensive two-dimensional gas chromatography (GCxGC) is a very powerful technique, based on the separation of gas phase analytes on two capillary columns with complementary interactions. A modulation device focuses the effluent from the first column in small segments, which are subsequently injected as narrow bands onto the second column, usually a short column suitable for high-speed separations.[77, 78] The use of two orthogonal columns results in a multiplicative increase in peak capacity in the GCxGC system in comparison to 1D-separations on the respective columns. In addition, the focusing step during modulation produces very narrow peaks and, consequently, lower detection limits.
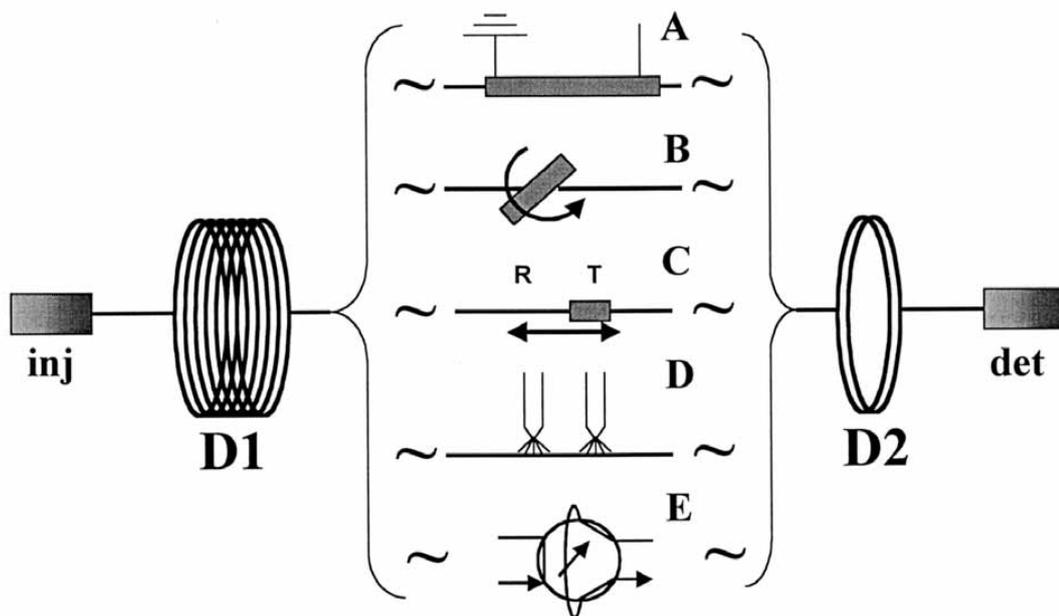
Figure 10: Schematic diagrams of various modulators used for GC×GC that trap solute coming from the primary dimension (D1) and pulse it to the second dimension (D2): (A) a heated tube encasing the capillary column; (B) the thermal sweeper system with a rotating slotted heater sweeping over the accumulation column to collect and pulse solute to the second column; (C) theLMCS that uses an oscillating cryotrap to trap the solute peak (position T), thus allowing it to heat up and remobilize by moving to the release position (position R); (D) a dual jet modulator using two jets to supply localized cryogenic cooling to the column; and (E) an example of the valve system that partially fills a loop and then switches the valve to flush the loop to the second column (the diagram shows the valve in the flush position). Taken from [78].

Data analysis is still a major problem for 2D separations, because programs available to date were derived from software designed for 1D separations. The analysis of known components can usually be done with proprietary software from instrument vendors such as HyperChrom (Thermo) and ChromaTOF (LECO). Parallel factor analysis (PARAFAC) can be used for the deconvolution and quantification of overlapping peaks in higher order data.[79] The newer PARAFAC2 algorithm is even able to handle less-than-ideal 2D data where for example run-to-run retention time shifting has occurred. However, one problem of this approach is the requirement for identical peak shapes for one particular analyte in all the analyzed runs.

GCxGC-TOF-MS was used for the analysis of the yeast metabolome upon cycling oxygen consumption.[79, 80] Applying several data analysis tools including PARAFAC and PCA, 44 metabolites of known identity, such as methyl citrate, myo-inositol, G6P and cystathionine, as well as 41 unidentified compounds were found to exhibit an oscillatory behavior over time.

Organic and amino acids of the central carbon metabolism were analyzed by GCxGC-TOF-MS in *M. extorquens* under two different growth conditions.[81] Quantification was only performed over one order of magnitude, by spiking target compounds into the samples. PARAFAC analysis was used to remove noise and overlapping peaks, and several changes in pool sizes were found that corresponded well with flux data.

GC-TOF-MS and GCxGC-TOF-MS were compared for the metabolomic analysis of mouse spleen tissue.[82] Seven times more abundant and high quality peaks were found with the two-dimensional analysis. However, the high-information content of the obtained chromatograms necessitated the use of sophisticated data analysis software. Some of these tools, such as direct chromatogram comparison, chromatogram subtraction and averaging routines as well as weighted chromatograms and Student's t-test, were compared and contrasted.[82]

4.3.6 Comparison of methods for the analysis of central carbon metabolites

The single quadrupole mass analyzer has the advantage of being inexpensive and robust. Used in SIM mode, it yields good detection limits and linearity. However, only unit mass resolution can be achieved and no MS/MS experiments can be performed. Triple quadrupole MS offers robustness and various MS/MS techniques. Using the instrument in MRM mode allows the specific and sensitive analysis of multiple analytes of interest. However, prior knowledge of the analytes is required for this scan type. Triple quadrupole MS is less sensitive in scan mode and gives only unit mass resolution, which leads to a high number of metabolite hits that cannot be discriminated. For

screening applications, a TOF-MS is more suitable. The new generation of TOF-MS has the advantage of yielding high mass accuracies (typically 3-5 ppm) at high data acquisition rates over a wide dynamic range. But since background ions cannot be removed, detection limits are higher than on a TQ operated in MRM mode, which, however, carries the disadvantage that only a limited number of metabolites can be measured. This may be overcome in the future with the introduction of scheduled MRMs (Applied Biosystems). Also, the TOF is more sensitive to environmental changes, thus necessitating regular recalibration and tuning.

CE-MS has the advantage of high peak capacities and resolution, while using only minute sample amounts. However, increasing injection amounts is difficult, therefore LODs are usually higher than those observed for LC-MS. Among LC-MS or LC-MS/MS methods, IP-LC-MS is well suited for polar analytes. Nevertheless, the IP reagent may cause ion suppression and fouling of the instrument. HILIC has the advantage of using solvents that are highly suitable for LC-MS analysis, yet retention times are often less reproducible compared to other LC methods. One advantage of HILIC over CE-MS is its ability to separate noncharged polar compounds such as sugars. While AEC provides reasonable resolution for negatively charged metabolites, it suffers from poor sensitivity, with LODs typically in the middle to high µM range, and a limited dynamic range. GC-MS, in comparison, is a very robust method and the excellent reproducibility of retention times and MS fragmentation patterns carries the advantage that analytes can be identified by mass spectral similarity and retention index comparison with commercial and custom libraries of standards. However, polar analytes are only amenable to GC-MS analysis after derivatization and thermolabile compounds cannot be measured. A comparison of important approaches towards metabolome analysis is shown in Table 1.

Table 1: Comparison of selected approaches for the metabolic analysis of the central carbon metabolism

Method	Advantages	Disadvantages	LOD	Ref.
MALDI-TOF-MS	• Well suited for screening • Simple sample preparation	• Poor linear range • Poor quantification • Poor reproducibility	15 nM – 1 µM	[42-46]
CE-TOF-MS	• No derivatization necessary • High peak capacities • Low sample consumption • High mass accuracy • Well-suited for screening	• High LODs reg. concentrations • Sensitive to sample ion strength • No fragmentation	0.2 – 2 µM	[49, 83]
PGC-LC-IT-MS	• High resolution of polar compounds • High resolution of isomers	• Often complicated gradients needed • Limited linear range	0.1 – 1.5 µM	[56-59]
UPLC-MS	• Increased efficiency • Higher peak capacities • Faster separations possible	• High pressure • Specialized equipment required		[61-63]
IP-LC-MS/MS	• Good resolution of polar compounds • Good linearity over a wide range • Very low detection limits • Very good reproducibility	• Ion suppression • Fouling of the MS caused by IP reagent	0.1 – 500 nM	[11, 66, 67]
HILIC-MS	• Good retention of polar compounds • Solvents well-suited for MS • Suited for uncharged compounds	• Poor reproducibility of retention times • Poor reproducibility		[3, 5, 69]
AEC-MS	• Well-suited for charged analytes • Good resolution of sugar phosphates	• Poor sensitivity • Limited dynamic range	42 – 440 µM	[70-73]
GC-MS	• Very robust method • High reproducibilities • Suitable for a wide range of analytes • Identification of unknowns via libraries	• Requires derivatization • Not suited for thermo labile compounds	40-500 pg on column	[74, 75]
GCxGC-TOF-MS	• Very high peak capacities • Improved S/N ratios	• Compound identification difficult • Quantitative capabilities have yet to be assessed		[77-82]

4.3.7 Flux analysis

The methods described above are mostly used for the analysis of the metabo-
lite composition of a given biological system. However, in order to investigate
the overall behavior of a network, it is important to also gain information on *in
vivo* reaction rates. Since it is not possible to measure these rates *per se*,
metabolic flux analysis uses different methods for the indirect measurement of
metabolic reaction rates.[84] Different strategies have been applied to analyze
and quantify intracellular fluxes and metabolic networks, analysis is usually
performed by mass spectrometry (MS) or nuclear magnetic resonance (NMR)
based methods.[85]

The most common approach consists of feeding the cells a labeled substrate,
usually ^{13}C labeled glucose, and then analyzing the label distribution, either in
proteinogenic amino acids or in metabolic intermediates. The analysis of
labeled amino acids is commonly done by GC-MS. For instance, this approach
was used to investigate different *E. coli* strains with mutations in the central
carbon metabolism. It was demonstrated that the oxidative PPP was substan-
tially decreased under anaerobic conditions in a wild type strain and that upon
a complete block of glycolysis the Entner-Doudoroff (ED) pathway contributed
about 30% to glucose catabolism.[86] In another study, flux analysis by GC-MS
led to the observation of substantial flux rerouting from glycolysis towards PPP
upon a block in the the direct transfer of a reducing equivalent from NADH to
NADPH.[12] Complex changes of carbon flux throughout the central metabolism
were observed in *S. cerevisiae* after a shift from oxidative to fermentative
growth.[87]

The second approach to metabolic flux analysis is the direct measurement of
labeling distributions in intermediates of the central carbon metabolism. The
analysis of flux distribution between glycolysis and PPP was performed by
GC-MS measurement of lactate and ribose labeling, derived from [1,2-
^{13}C$_2$]glucose.[88, 89] Isotopomers of intracellular metabolites in *E. coli* were also
measured by CE-TOF-MS.[90] The results were compared with data from GC-

MS analysis of proteinogenic amino acids. The correlation between the two methods was quite high, remaining differences were contributed to alterations in labeling distribution between precursors and amino acids as well as influences from flux modeling.[90]

Several effects have to be taken into account when performing flux analysis by GC-MS. The sample has to be prepared by chemical derivatization, as described above, mainly by oximation and silylation.[12, 86] However, silicon has a significant percentage of naturally occurring isotopes, namely 4.7% ^{29}Si and 3.09% ^{30}Si. Also, strong isotope effects have been observed, which led to different retention times of differently labeled isotopomers in GC columns.[91]

There are only few reports on the use of LC-MS based methods in the context of flux analysis. The advantage of these methods is that no derivatization is necessary, thus correcting for natural isotopes is easier. An IP-LC-MS/MS based method was used to study the effect of a short glucose pulse in *E.coli* applying ultra-fast sampling.[92] Anion-exchange chromatography was used to evaluate mass isotopomer distributions of primary metabolites, which was subsequently fitted with a model to yield metabolic flux data.[93] A method for flux analysis in fed batch cultures by measuring intracellular free amino acids was described.[94] However, amino acids had to be derivatized prior to LC-MS/MS in order to increase their ionization efficiency.

4.3.8 Platform approaches

The future of metabolomics is definitely its integration with genomics and proteomics in order to gain a complete picture on cellular processes. The first step in this regard will be to combine different analytical techniques for a comprehensive coverage of the metabolome. For example, UPLC-MS, HILIC-MS and GC-TOF-MS approaches were combined for the analysis of the urinary metabolome.[95] Automated peak picking was performed with MZMine and XCMS, both freely available online. The extracted features were then used to discriminate samples from patients with renal cell carcinoma from healthy controls (n=6 each). For all three methods significant features were detected

that could discriminate the two groups, however no identification of these features was attempted.

A more comprehensive approach employed six different methods for metabolome analysis for microbial metabolomics.[96] First, available metabolome information from three microorganisms (*E. coli*, *B. subtilis* and *S. cerevisiae*) was compiled in a list that comprised 905 different metabolites. The list was generated from the metabolite databases for these organisms after manual curation. Then, three GC-MS and three LC-MS methods were used to analyze 399 compounds commercially available, 380 of which were ultimately amenable to analysis. When applied to exponentially growing *E.coli* cell extracts, 431 peaks were detected and 176 metabolites could be identified, including 61 that were not previously annotated in existing *E. coli* databases.

The integration of multiple high-throughput measurements of alterations in mRNA, proteins, metabolites and fluxes was recently done by Ishii et al. [97] An *E. coli* wild type strain and 24 single-gene mutants were analyzed by six different methods. Gene expression was measured by DNA microarrays and qRT-PCR, protein expression by two-dimensional differential gel electrophoresis and LC-MS/MS, metabolite levels by CE-TOF-MS and fluxes by GC-MS of ^{13}C-label distributions in proteinogenic amino acids. The authors could demonstrate that the intracellular metabolic network in *E. coli* was stable towards various types of perturbations.

4.4 Sample preparation

Sample preparation is of high, but often neglected importance for the general validity of results obtained from a metabolomic experiment. The two most crucial steps in sample pretreatment are the rapid quenching of metabolism and the extraction of metabolites. Both steps may introduce systematic or random errors. For example, if the effect of a pulsed perturbation on the metabolism is investigated, fast sampling and quenching of the metabolism is

of utmost importance, since numerous metabolites have turnover times of less than one second.[57]

However, the most common quenching method, which employs rapid transfer of the sample into a cold methanol:water mixture, leads to substantial metabolite leakage from cells that do not possess a strong cell wall, such as different bacterial strains including *E. coli* or human cell lines.[98-100] Alternative methods for quenching have been described, for example the fast sampling into a syringe containing pre-cooled stainless steel beads.[100] Very rapid sampling with intervals of less than 5 s and immediate quenching on a subsecond time scale (<200 ms) was described for *S. cerevisiae*.[73] Another strategy is the combined measurement of intra- and extracellular metabolites, for example after rapid heating of the sample to 95ºC, where intracellular metabolites are released quantitatively to the surrounding medium.[101] However, for this approach extracellular metabolite concentrations have to be substracted, which increases the influence of measurement errors. A method for steady-state measurements, where rapid changes are not supposed to occur, is the fast filtration of the broth, to separate cells from the media. This approach was found to substantially reduce leakage, especially for gram-positive strains.[98, 99]

Secondly, the metabolites have to be released from the cells and this process should be complete or at least reproducible. A number of different extraction methods were proposed,[102] but only a few studies have compared different methods to date.[75, 103, 104] It is to be expected, that different methods will yield optimal results for different organisms or metabolite classes. This was comprehensively demonstrated in a recent evaluation of different solvent mixtures for the extraction of metabolites from *E. coli*.[105] It was found that 80% methanol/water is well suited for many metabolites. However, a number of phosphorylated compounds, especially nucleotide triphosphates, are severely reduced under these conditions. For these compounds, mixtures of acidic acetonitrile/water (80:20) or acetonitrile/methanol/water (40:40:20) gave superior yields due to reduced decomposition. Optimization of extraction protocols for mammalian cells was performed by fractional factorial design,

based on the evaluation of different conditions on four response variables.[106] Methanol/chloroform and boiling methanol yielded overall the best results.

4.5 Data analysis

4.5.1 Handling of complex metabolomic datasets

Typical metabolomic experiments can generate large amounts of data. Handling such complex datasets has a huge impact on the quality of metabolite identification and quantification, and thus on the ultimate biological interpretation of the results. This has led to an increased development of methods and software tools for metabolomics in the recent years. The common data processing routine consists of several stages, including filtering, feature detection, alignment and normalization.[107] However, in order to process the raw data derived from different proprietary mass spectrometer softwares, it has to be converted into a common data format, usually netCDF or mzXML. Several programs are available for feature detection [107], both commercial, e.g. MarkerLynx, MassHunter, Metabolic Profiler, metAlign, and freely available, e.g. MathDAMP, MET-IDEA, MZmine, XCMS. MZmine [108, 109] has the advantage that it can handle both netCDF and mzXML data and align the extracted features from multiple samples. While MZmine (Figure 11) is well suited for CE-MS and LC-MS analysis, it does not perform spectral reconstruction. This complicates data-sets derived by EI-GC-MS, where considerable fragment generation occurs. The program MET-IDEA [110] is able to use data extracted from AMDIS, a software used for the deconvolution of GC-MS data. MET-IDEA eliminates noise, corrects for retention time shifts and automatically detects and quantifies all potential metabolites.
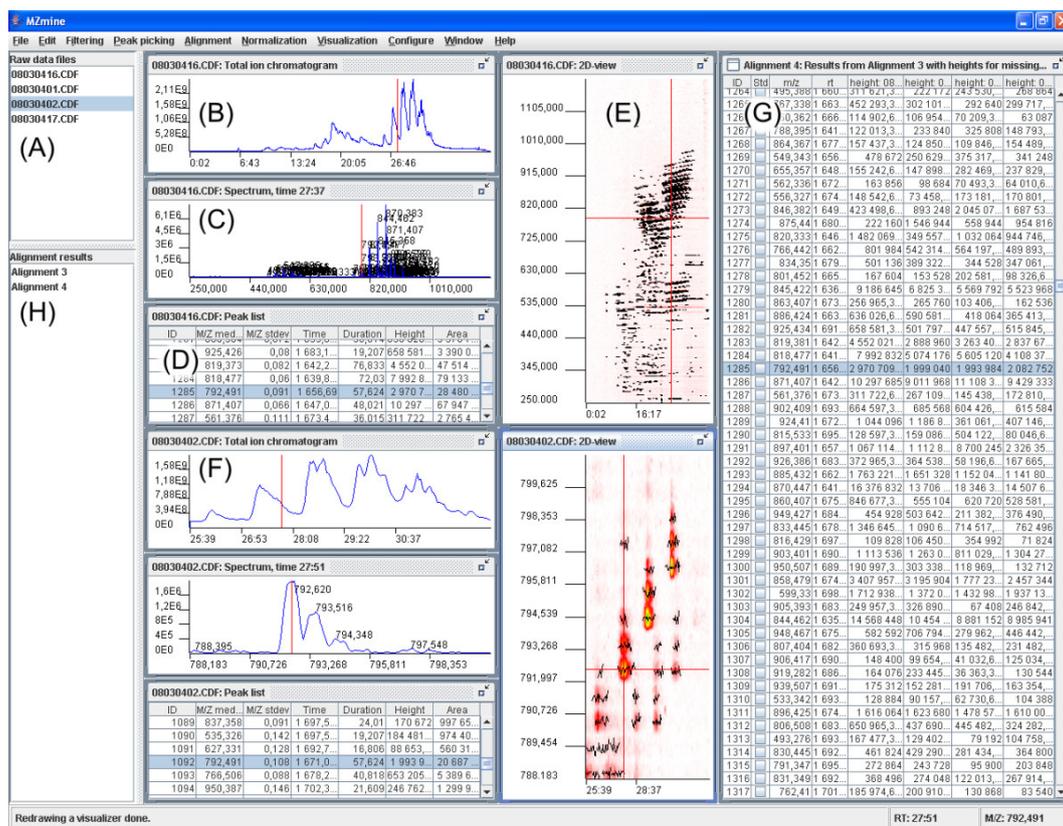
Figure 11: *MZmine* graphical user interface: (A) List of imported raw data files. (B) Total ion chromatogram (TIC) for selected file. (C) Mass spectrum for selected retention time for the same file. (D) Peak list for the same file, with listed m/z values, retention times, and intensities. (E) 2D map of the same file, with retention time on x-axis and m/z on y-axis. (F) Zoomed-in spectra for a different file. (G) Peak alignment matrix for all files listed. (H) Available alignment results, e.g. for different normalizations. Taken from [109].

The extracted features from CE- or LC-TOF-MS data are usually characterized by their exact mass, their retention / migration time and potentially their molecular formula. Especially the mass information can now be used to query a database for potential metabolite identification. A comprehensive collection of available databases can be found in [111, 112]. In order to confirm the obtained hits, ideally a commercial standard can be measured. If this is not possible, the structure of the putative identification can be cross-checked with closely eluting compounds to check for plausibility. If available, the molecular formula is also a valuable confirmation.

In GC-MS is the identification of unknown metabolites poses additional problems due to the fragmentation occurring under electron impact ionization.

Usually identification is performed by database search, commonly the AMDIS database, however if a compound is not included in the reference database, it can not be identified and tracked with this approach. Recently, a program termed SpectConnect was introduced, which compares every spectrum in each sample to the spectra in every other sample.[113] By this it can determine which components are conserved across replica samples and also detect differences between different conditions.

4.5.2 Quantification

Quantification is an issue of high importance in metabolomic studies. Knowledge on absolute quantities of metabolites is important for the modeling of cellular networks and to evaluate responses to external stimuli. Initially, quantitation was often performed in a relative fashion or by using a single internal standard compound as reference. Recently, the use of several internal standard compounds, often stable isotopes of key analytes, was shown to yield better reproducibilites, especially if standards are introduced at different stages during sample preparation and measurement.[74, 83, 114] An elegant method to account for many different potential interferences, such as partial degradation of metabolites, non-linearity, instrument drift or matrix effects, was introduced recently.[115] *S. cerevisiae* was grown on [U-^{13}C]glucose, the extracted metabolites were subsequently used as internal standards for metabolite quantification. The method was termed "Mass Isotopomer Ratio Analysis of U-^{13}C Labeled Extracts" (MIRACLE).[116] The same mixture was used later for the analysis of metabolites of *Penicillium chrysogenum*. Intracellular metabolites were quantified with relative standard deviations (RSDs) of ≤5%.[117]

5 Experimental section – Materials and Instrumentation

5.1 Chemicals

All metabolite standards, ammonium acetate, LC-MS grade water, 2-propanol (LC-MS grade), 1,2-bis(3-aminopropylamino)ethane, epichlorohydrine, LB and M9 broth, methoxylamine hydrochloride, amino acid standard solution, pyridine, LC-MS grade acetic acid, tributylamine and [U-^{13}C, U-^{15}N]AMP were from Sigma (Sigma-Aldrich, Taufkirchen, Germany). Methanol (LC-MS) was from Fisher (Fisher Scientific GmbH, Ulm, Germany). All solvents were HPLC-MS grade. [U-^{13}C]lactate, [U-^{2}H]succinate and [2,2,4,4,4-^{2}H$_5$]glutamic acid, [U-^{13}C]glucose, [2,3,3,3-^{2}H$_4$]lactate , [U-^{13}C]fumarate and the [U-^{13}C, U-^{15}N] cell free amino acids mixture were from Euriso-top (Saint-Aubin Cedex, France). DL-[2,3,3-^{2}H$_3$]malic acid and [U-^{13}C]pyruvate (MSTFA) were purchased from C/D/N Isotopes Inc. (Quebec, Canada). *N*-methyl-*N*-trifluoroacetamide was from Macherey-Nagel (Dueren, Germany). [U-^{13}C]glucose-1-phosphate and [1,2-^{13}C$_2$]glucose were from Omicron Biochemicals Inc. (South Bend, IN, USA). The EZ:faast kit for GC-MS was obtained from Phenomenex Inc., Torrence, CA, USA.

5.2 Instrumentation

Capillary electrophoresis was performed on an HP3D-CE system (Agilent Technologies, Palo Alto, CA, USA). For CE-TOF-MS the CE capillary was coupled to a MicrOTOF Focus (Bruker Daltonics, Bremen, Germany) via a coaxial CE-ESI sprayer. Sheath liquid was delivered by an 1100 series isocratic pump (Agilent Technologies).

Shake flasks for cultivation of *E. coli* cells were incubated in an Innova 4430 Incubator Shaker (New Brunswick Scientific GmbH, Nuertingen, Germany).

Optical density measurements were performed on an Ultrospec 3100 pro UV/VIS Spectrophotometer (Amersham Biosciences, Buckinghamshire, UK). Solvent evaporation during sample preparation was performed on a Combi Dancer Infra-Red Vortex-Evaporator (Hettich AG, Baech, Switzerland).

A pH-electrode Lab 850 (SCHOTT Instruments GmbH, Mainz, Germany) was used for the adjustment of pH values of buffers and solvents.

For the sterilization of all the materials used for the cultivation and preparation of bacterial and cell culture samples a steam sterilizer Varioklav (Thermo Fisher Scientific Inc, Waltham, MA, USA) was used.

GC-MS for most metabolites was performed on an Agilent Technologies Model 6890 GC equipped with a 5975 Inert XL Mass Selective Detector (MSD) and a 7683B Auto Liquid Injector. GC-MS analysis of amino acids was performed on an Agilent Technologies Model 6890 GC equipped with an 5975 Inert XL MSD, a PTV injector (Gerstel, Muehlheim, Germany) and a MPS-2 sample robot (Gerstel).

An Agilent 1200 SL HPLC system (Agilent Technologies, Waldbronn, Germany) was used for liquid chromatography. The HPLC system was directly coupled to a 4000 QTRAP mass spectrometer (ABI/MDS Sciex, Concord, Canada). Direct infusion was performed using a model 11 PLUS syringe pump (Harvard Apparatus, Holliston, USA).

5.3 Software

Analysis of the data generated on the MicrOTOF instrument was performed using DataAnalysis 3.4 and QuantAnalysis 1.8 (Bruker Daltonics, Bremen, Germany). Peak picking and alignment were performed with MZMine, version 0.60. Analysis of data generated on the 4000 QTRAP instrument was performed with Analyst 1.4.2.

6 CE-TOF MS method development

6.1 Introduction

As described in chapter 2, different methods are used for metabolome analysis. The feasibility of CE-MS for metabolomics and its quantitative capabilities have been demonstrated in several studies.[10, 25, 50, 118-122] Recently, the coupling of CE to a TOF-MS was shown to yield much more information in a single run, than could be achieved on a single quadrupole instrument.[49]

Our goal was the development of a CE-TOF-MS method for quantitative metabolic profiling of negatively charged metabolites. To this end, we used capillaries coated with a cationic polymer to reverse the EOF. We optimized the coating procedure as well as the background electrolyte and the sheath liquid. A method to correct for shifts in migration time was introduced. Finally the quantitative capabilities of the method for 20 selected metabolites involved in the central carbon metabolism were established and several quality control steps were introduced to ensure a reliable measurement over an extended period of time. A shortened version of this chapter was published as part of [83].

6.2 Methods

6.2.1 Experimental details of the optimized CE-TOF-MS method

Capillary electrophoresis was performed on an HP3D-CE system (Agilent Technologies, Palo Alto, CA, USA). Prior to use, fused-silica capillaries (Polymicro Technologies, LLC, Phoenix, AZ, USA; 100 cm x 50 µm ID) were coated with PolyE-323 for 20 min as described in [123] to reverse the EOF. After coating, the capillaries were equilibrated for 20 min with background electrolyte (BGE), containing 50 mM ammonium acetate, pH 8.7, in 5% (v/v) methanol in water. Sample was injected by applying a pressure of 50 mbar for 30 s (25.8 nL, 1.31% of capillary volume). Metabolites were separated by applying a voltage of -30 kV for 21 min. The capillary was washed between

runs with 10 mM sodium hydroxide and water, and then re-equilibrated with BGE for 7 min. Total analysis time amounted to ~30 min.

The capillary was coupled directly to a MicrOTOF Focus (Bruker Daltonics, Bremen, Germany). A 50% (v/v) isopropanol:water solution containing 20 mM ammonium hydroxide served as sheath liquid in the coaxial CE-ESI sprayer, delivered by an 1100 series isocratic pump (Agilent Technologies) at 4 µL/min. Ionization was performed in negative mode at 4500 V, assisted by a nebulizer pressure of 0.4 bar and a drying gas flow of 6 L/min at 200°C. The TOF detector was set to scan a mass range of 40-800 m/z at 3 scans/s. Resolution of the TOF-MS was typically ~10,000 at m/z 100 and ~15,000 at m/z 300. Mass recalibration was performed during post-processing by using sodium acetate clusters formed during the wash step with sodium hydroxide, leading to a reproducible mass accuracy of ~3-5 ppm.

Absolute quantitation of 20 different metabolites, including 2KG, F6P, G1P, G6P, FBP, AMP, ADP, NAD, glutamate, aspartate, glycerol-1P, glycerol-2P, fumarate, oxaloacetate, lactate, pyruvate, malate, succinate, citrate, and 3PG, was accomplished with a dilution series of an equimolar (100 mM) aqueous stock solution of all standards (= standard mix), to which 100 µM each of the extraction and internal standard compounds [U-^{13}C]lactate, [U-^{2}H]succinate, glutamic acid-d$_5$, malic acid-d$_3$ and PIPES had been added. Prior to analysis of biological samples, 11 calibration points were generated over a concentration range of 1-1000 µM. Spectral peak integration was done with QuantAnalysis (Bruker Daltonics) following normalization of the peak areas of the standards to the area of the closest migrating internal standard. Concentrations were then inferred from the calibration curves. Stability of measurement was checked by periodical injection of a quality control sample between biological samples.

6.3 Results and Discussion

6.3.1 Choice of capillary coating

At pH-values above 5, the negatively charged silanol groups on the inner surface of a bare fused silica capillary generate an EOF towards the cathode during electrophoresis, which is unfavorable for the analysis of anions. This may be remedied by coating the inner capillary surface with a cationic layer, resulting in the reversal of the EOF. Initially, capillaries commercially coated with hexadimethrine bromide [10] were used. However, in our hands, they yielded poor reproducibility of migration times and quantitative measurements, respectively (data not shown). Therefore, we switched to PolyE-323, a cationic polyamine.[123] Preparation of the PolyE-323 coating solution proved simple and inexpensive. Coating was performed off-line with a PHD 2000 syringe pump from Harvard Apparatus (Holliston, MA, USA) to prevent fouling of the ion source with the cationic polymer. The coating remains stable for at least 60 runs. A drop in CE current of >20% indicates that the capillary has to be re-coated or a new capillary has to be prepared.

6.3.2 Optimization of the background electrolyte

Ammonium acetate was used as BGE, because its ions are volatile and do not interfere with mass spectrometry. Of the different concentrations of ammonium acetate evaluated (25 mM, 40 mM, 50 mM, and 60 mM), 50 mM gave baseline resolution for all compounds except F6P and G6P (R_s~1.3), see Figure 12 for a chromatogram under optimized conditions. In addition, a reasonably fast analysis time of ~30 min per sample, including pre- and post-conditioning, was obtained. Using buffers with lower ionic strength significantly increased analysis time, while an ammonium acetate concentration of 60 mM failed to resolve F6P, G6P and G1P.

Next, different pH values were tested (8.5-9.1). Optimization parameters were resolution of the hexose-phosphates, relative migration times (MT) of the metabolites, peak shape and coating stability. The coating was stable under

all pH conditions, but peaks substantially broadened at pH values >9. A pH of 8.7 was found to yield overall the best results.

We also tested the addition of different organic modifiers to the BGE.[118] A methanol concentration of 5% (v/v), to reduce unwanted analyte-wall interactions, in 50 mM aqueous ammonium acetate, pH 8.7, gave the best results for the resolution of the hexose-phosphates, while peaks broadened significantly at methanol concentrations >10%. The addition of acetonitrile or isopropanol to the BGE yielded electrophoretic separations inferior to methanol.

6.3.3 Sheath liquid optimization

The use of isopropanol (50% v/v in water) as sheath liquid proved superior to methanol in ensuring spray stability. As the volatile ionic constituent in the sheath liquid we chose ammonium hydroxide, because it increases the pH of the solvent and, thus, increases ionization efficiency in negative mode. We tested different concentrations of ammonium hydroxide (5 mM, 20 mM, 50 mM and 280 mM). Performance was best at 20 mM. At 5 mM, the ionization efficiency of certain analytes (*e.g.,* aspartate, citrate, fumarate) deteriorated. Ionization did not increase at concentrations >20 mM, but the mechanical stability of the fused silica capillary was affected. Especially at the highest concentration, the capillary became brittle and the polyimide coating detached from the outer capillary surface over the course of a day. Therefore, 20 mM ammonium hydroxide in water-isopropanol (50:50, v/v) was used for all CE-MS analyses as the sheath liquid, which was delivered by an isocratic pump at a flow-rate of 4 µL/min. The extracted ion chromatograms (EIC) for selected metabolites under optimized method parameters are shown in Figure 12.
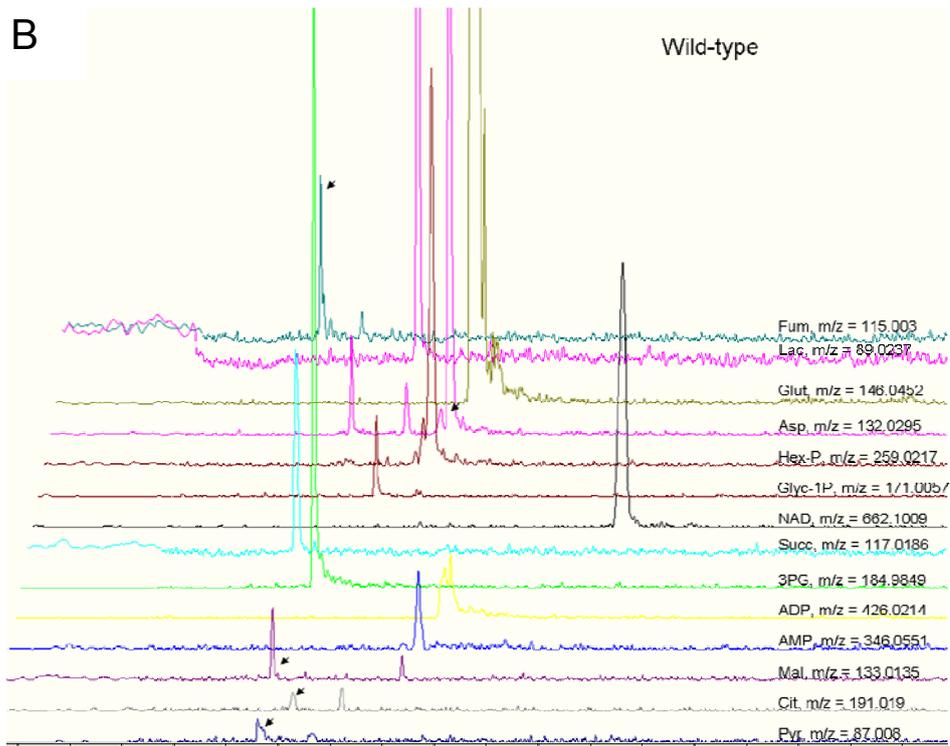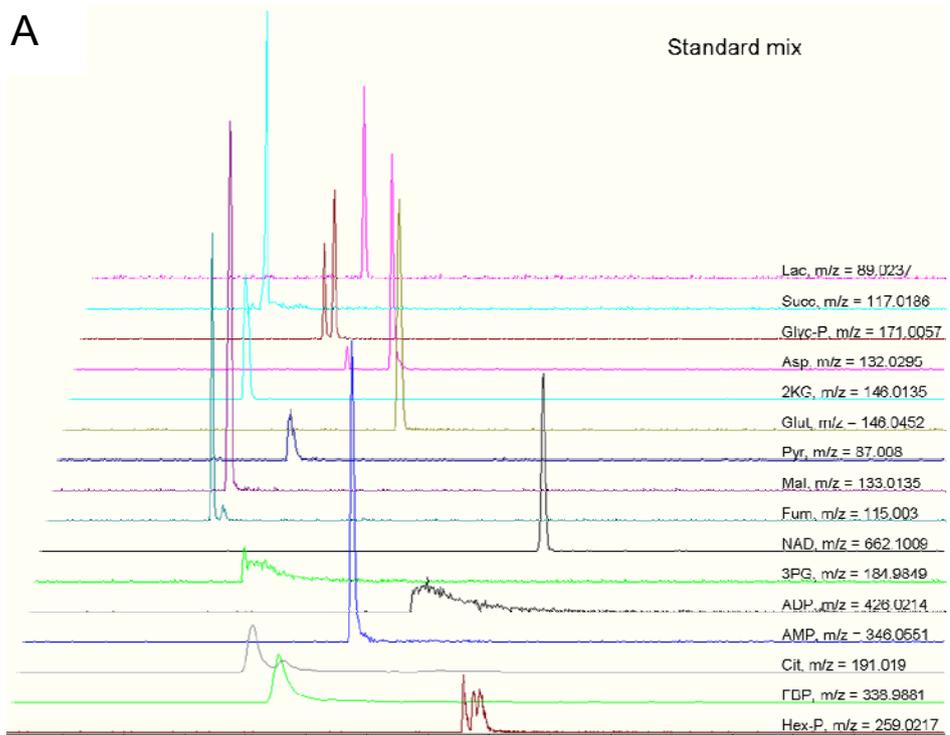
Figure 12: Extracted ion chromatograms of A) a mixture of standards (31.25 µM each) and B) an *E.coli* wild type sample (grown in M9 medium). Lac, lactate; Succ, succinate; Glyc-P, glycerol-phosphate; Asp, aspartate; 2KG, 2-ketoglutarate; Glut, glutamate; Pyr, pyruvate; Mal, malate; Fum, fumarate; Cit, citrate; Hex-P, hexose-phosphate.

6.3.4 Migration time shift

A common problem in CE is MT shift between injections. This effect is dependent on the ionic strength of a sample, which leads to different stacking conditions and, thus, to different effective field strengths in the capillary. In addition, the complex matrices of biological samples often contain components that interact with the inner capillary surface, leading to alterations in EOF.[124, 125] To minimize the latter, capillaries were flushed sequentially with BGE (30 s), 10 mM NaOH (1 min) and water (30 s), followed by equilibration with BGE for 7 min. Nevertheless, MT was still not sufficiently constant to serve as a reliable criterion for metabolite identification. Therefore, we implemented an algorithm that corrects for MT shifts as proposed by Reijenga et al.[49, 124] Whereas the latter used two arbitrary sample peaks, we chose to use internal standards, thus eliminating the need for prior knowledge of analyte composition. To this end, three internal standards, namely glutamic acid-d$_5$, malic acid-d$_3$ and PIPES, were added to every sample at a final concentration of 100 µM. Standard runs in freshly coated capillaries were used to determine migration times under reference conditions.
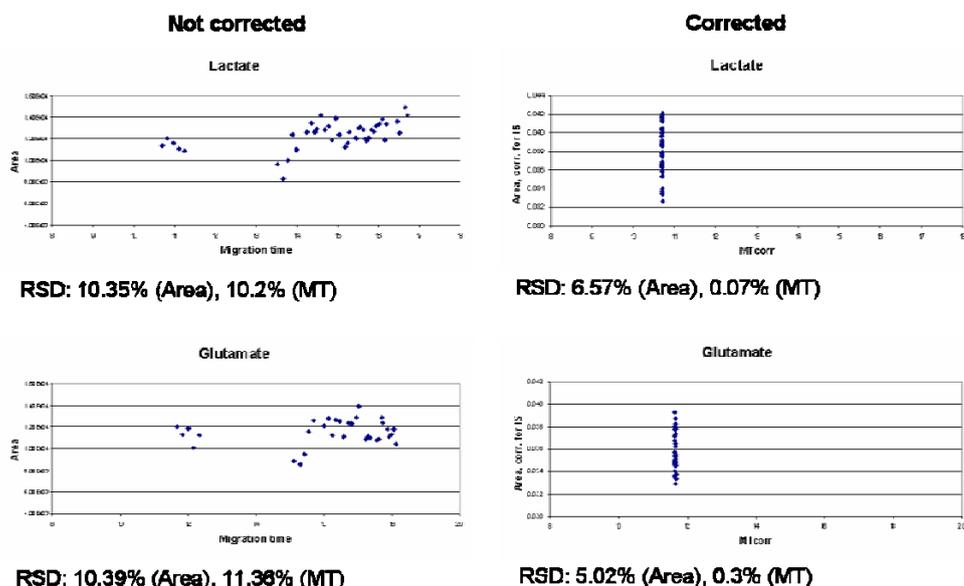
Figure 13: Effect of correcting for spray fluctuations and MT shift. Glutamic acid-d5, malic acid-d3 and PIPES are used as internal standards, n=55; left: not corrected, right: corrected.

We tested the correction method by measuring a total of 45 injections of metabolite standard and 5 injections each of two different biological samples in one batch. Relative standard deviations for peak areas were ~10-15%, for migration times ~10%. They were reduced to 5-7% for peak area and 0.07-0.13% for migration time after applying the corrections described above (Figure 13). This facilitates subsequent spectral alignment and, thus, identification of analytes.

6.3.5 Quantitative metabolite analysis and quality control

In order to achieve absolute quantitation of selected metabolites, calibration curves were generated. Calibration curve parameters as well as LOD and limits of quantitation (LLOQ, ULOQ) are presented in Table 2.

Table 2: Calibration curve parameters for selected metabolites

Metabolite	RSQ[a]	LOD[b] (µM)	LLOQ (µM)	ULOQ (µM)
Lactate	0.9997	1	2	5000
Pyruvate	0.9985	1	4	5000
Oxaloacetate	0.9987	5	30	5000
Citrate	0.9985	1	4	5000
2KG	0.9986	0.5	1	2500
Succinate	0.9975	0.5	1	2500
Fumarate	0.9999	0.5	2	2500
Malate	0.9963	0.25	1	5000
Aspartate	0.9988	0.5	1	2500
Glutamate	0.9998	0.5	1	5000
G1P	0.9983	0.5	1	500
G6P	0.9997	1	2	500
F6P	0.9975	1	2	500
FBP	0.9992	2	4	2000
3PG	0.9962	2	4	2000
Glycerol-1P	0.9966	1	2	2500
Glycerol-2P	0.9988	1	2	2500
AMP	0.9997	0.5	1	2500
ADP	0.9999	2	4	2500
NAD	0.9993	0.5	2	2500

a) Square of the correlation coefficient r of the regression analysis
b) Defined as a S/N ≥ 3

The LOD was defined as the concentration, at which a compound could be detected with a signal-to-noise (S/N) of ≥3. The LLOQ and ULOQ were defined according to the FDA Guide for Bioanalytical Method Validation [126] as the lowest and highest points of the calibration curve, respectively, at which a compound could be measured with an accuracy between 80-120%.

A number of quality control (QC) measures were introduced to ensure reliable and accurate quantitation of the metabolites of interest. Alterations of metabolite levels during sample preparation were accounted for by adding 100 µM each of $[U-^{13}C_3]$lactic acid, $[U-^{13}C_6]$glucose, norvaline, succinic acid-d_4 and cinnamic acid-d_7 in methanol to every sample immediately before extraction. The extracts were dried by evaporation and reconstituted in MilliQ-water containing 100 µM each of glutamic acid-d_5, malic acid-d_3 and PIPES. These internal standards were used to correct for shifts in migration time and alterations in electrospray efficiency (as described above, see Figure 13). Samples within a batch were run in a random order. Prior to each batch, 11 calibration standards of different concentrations were analyzed to check for calibration curve stability. After every 8-10 biological samples, several QC samples were analyzed, consisting of 3 calibration standards (low, medium and high concentration), one biological replicate and a blank sample containing only internal standards. The QC samples and the samples associated with them were accepted only if at least 67% of the QC samples were within ±15% of their nominal value.[126] The intra-day analytical variability of metabolite standards was 0.7 – 12.5%, while the inter-day variability ranged from 4.1% to 12.3% (Table 3).

Table 3: Reproducibility of CE-TOF-MS and extraction of selected metabolites, and biological variability of four replicates each for two *E. coli* strains.

Metabolite	Variance, Std (%)	Variance, *E. coli* samples (%)			
	Intra[a]- / Interday[b] n=4 / n=8	Analytical[c] n=4	Extraction[d] n=3	Biological[e] P-U, n=4	Biological[e] wt, n=4
Lactate	0.7 / 4.5	12.4	16.5	23.9	19.5
Pyruvate	6.7 / 9.1	5.4	7.9	24.8	19.1
Citrate	10.4 / 11.4	9.6	11.3	4.5	14.0
2KG	10.6 / 8.4	9.2	13.7	10.6	17.1
Succinate	6.2 / 6.0	5.1	9.1	11.5	13.0
Fumarate	11.4 / 10.7	4.1	10.6	19.8	27.6
Malate	3.6 / 4.9	8.3	9.9	14.8	5.0
Aspartate	8.5 / 8.6	7.8	9.0	7.3	14.3
Glutamate	4.8 / 5.3	6.1	11.7	3.4	9.7
G1P	2.7 / 8.9	10.6	14.7	13.0	9.7
G6P	6.0 / 4.1	22.7	22.6	15.7	31.3
F6P	5.1 / 10.4	19.3	15.4	18.2	31.3
FBP	9.9 / 11.2	12.3	13.6	24.2	22.9
3PG	2.9 / 8.5	12	11.7	18.0	22.3
Gycerol-1P	12.5 / 8.9	10.8	15.6	17.4	14.9
AMP	4.8 / 7.0	3.9	11.9	30.7	5.1
ADP	8.9 / 9.8	12.6	17.0	29.2	18.1
NAD	11.2 / 12.3	3.7	11.4	10.4	24.8

7 Optimization of sampling and method validation for *Escherichia coli*

7.1 Introduction

Sample preparation is of high importance for the validity of results obtained from a metabolomic experiment. The two most important points in this regard are the rapid quenching of metabolic conversions after sampling and the sample work-up aimed at making all the metabolites amenable to analysis. Both steps have a high potential of introducing systematic or random errors if they are not sufficiently investigated and standardized and they have to be adapted to the specific biological problem at hand. Examples of different approaches for quenching and extractions were discussed in chapter 4.4.

Considering all these factors we decided to use fast filtration for sample harvesting. We evaluated the filtration process with regard to different filter materials, as well as several extraction methods. We found significant differences for different sample preparation methods. In addition, we performed a matrix spike experiment to establish the effect of the biological matrix on the absolute quantification of a selected set of metabolites. The results presented in this chapter were published as part of [83].

7.2 Material and Methods

7.2.1 Bacterial strains and growth conditions

The *E. coli* wild type strain MG1655 and the mutant UdhA-PntAB (MG1655 *ΔudhA ΔpntAB*) were cultured at 37ºC in 250 mL shake flasks containing 30 mL of Luria-Bertani (LB) medium. After inoculation with a preculture in the same medium to an OD of 0.1, cells were harvested in stationary phase. Cell dry weight (cdw) was determined by vacuum filtration of 10 ml medium

containing cells at different optical densities. The filters were dried at 120ºC and weighed empty and with dried cells.

7.2.2 Experimental details for cell harvesting and metabolite extraction

2-10 mL of cell suspension (volume depending on OD) and pure medium (for method blanks and filter material evaluation), respectively, were filtrated by vacuum filtration using polyethersulfonate (PESU) filters resp. cellulose acetate, cellulose nitrate or regenerated cellulose for comparison of filter materials (0.45-µm pore size, 25 mm, Sartorius AG, Goettingen, Germany). The filters were then washed with an identical volume of NaCl solution (0.9%, room temperature), before they were submersed in 3 mL of quenching/extraction solution in a 50 mL Falcon tube (Fisher Scientific, Pittsburgh, PA, USA). The extraction solution consisted of 80% (v/v) methanol in water at -20ºC, containing the extraction standards [U-^{13}C]lactate, [U-^{2}H]succinate, [U-^{13}C]glucose, cinnamic acid-d7, and norvaline. The tube was vortexed for 45 s, the filter was removed and checked visually for complete dissolution of cells.

The Falcon tube was transferred to liquid nitrogen for 3 min, thawed in an ice bath for 10 min, and briefly vortexed. This freeze-thaw cycle was repeated three times for complete cell disruption. The sample was centrifuged at 4ºC and 4.500 rpm for 8 min. The supernatant was transferred to a new tube. The pellet was re-extracted twice with 0.5 mL of 80% (v/v) methanol at -20ºC. The combined extracts were split into two aliquots, one for CE-TOF-MS and GC-(AA)-MS analysis, the other for GC-MS analysis. The samples were dried for 4 h in a vacuum evaporator (CombiDancer, Hettich AG, Bäch, Switzerland) and stored at -80ºC. For CE-MS analysis, the dried samples were dissolved in 100 µL of internal standard solution containing 100 µM each of glutamic acid-d$_5$, malic acid-d$_3$, and PIPES.

7.2.3 CE-TOF-MS

See Capter 6.2.1, optimized method parameters

7.3 Results and Discussion

7.3.1 Optimization of cell harvesting and metabolite extraction

The rapid quenching of the cellular metabolism is critical in capturing the immediate metabolic state and not an artificial distribution of metabolites generated by the sample handling process. The method commonly used is quenching in cold methanol. However, several studies found this procedure to cause a substantial leakage of metabolites, especially from prokaryotic cells such as *E. coli*.[98, 99, 104] Instead we chose to harvest cells by filtration, which can be performed faster than centrifugation.[99, 104] It takes about 1 min to filtrate and wash the sample. The filters were then immediately plunged into the cold extraction solution. This method might not be suitable for sampling continuous cultures or for evaluating pulse experiments, as some metabolites have turnover times of about 1 s. However, initially we were only interested in the metabolic profiles of batch cultures in stationary phase, which we expected to contain a high number of different metabolites.

Care has to be taken with regard to the filter material used for separating cells from medium. We tested four different filter types (cellulose acetate - CA, cellulose nitrate - CN, polyethersulfone - PESU and regenerated cellulose - RC) by filtering pure medium, followed by a wash with 0.9% NaCl solution and extraction with 80% methanol as described below. Lactate, glutamate, aspartate, malate, succinate, glycerol-2P, and glycerol-1P were quantitatively determined in the extracts (Figure 14). We found that CN adsorbed most metabolites to a high extent, while the other three materials showed only minimal absorption. The lowest adsorption was seen for PESU, which was then chosen for further experiments.
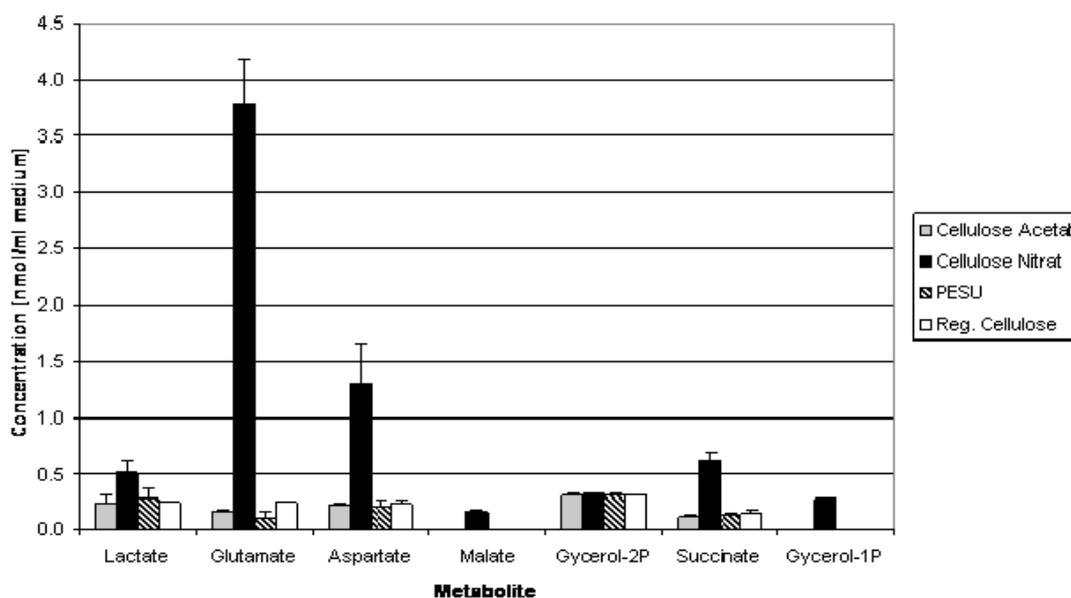
Figure 14: Unspecific adsorption of metabolites to different filter materials. Ten mL of fresh LB medium (without cells) was filtered in triplicate through 4 different filter materials. The filters were washed with 0.9% NaCl, before extraction and analysis as described under 7.2.2.

A number of methods have been reported for the extraction of metabolites from biological matrices. To date, however, only a few studies have compared directly the different extraction methods.[75, 104] We adapted the cold methanol protocol by Maharjan et al. [104] that reportedly gave the best results for *E.coli*. We tested in triplicate 50%, 80% and 100% methanol for extraction efficiency of metabolites from *E. coli* (Figure 15). The extraction yield of NAD was significantly higher in 80% than in 50% methanol. For carboxylic acids other than pyruvate, which could not be extracted in 100% methanol, no significant differences in extraction yield were observed between the three methanol concentrations tested. Contrary to findings in yeast [75], the use of 100% methanol resulted in lower extraction yields of glutamate and aspartate. Finally, in concordance with Kimball and Rabinowitz [103], we found 80% methanol to work best for the extraction of phosphorylated metabolites, which was chosen for all subsequent experiments.

55
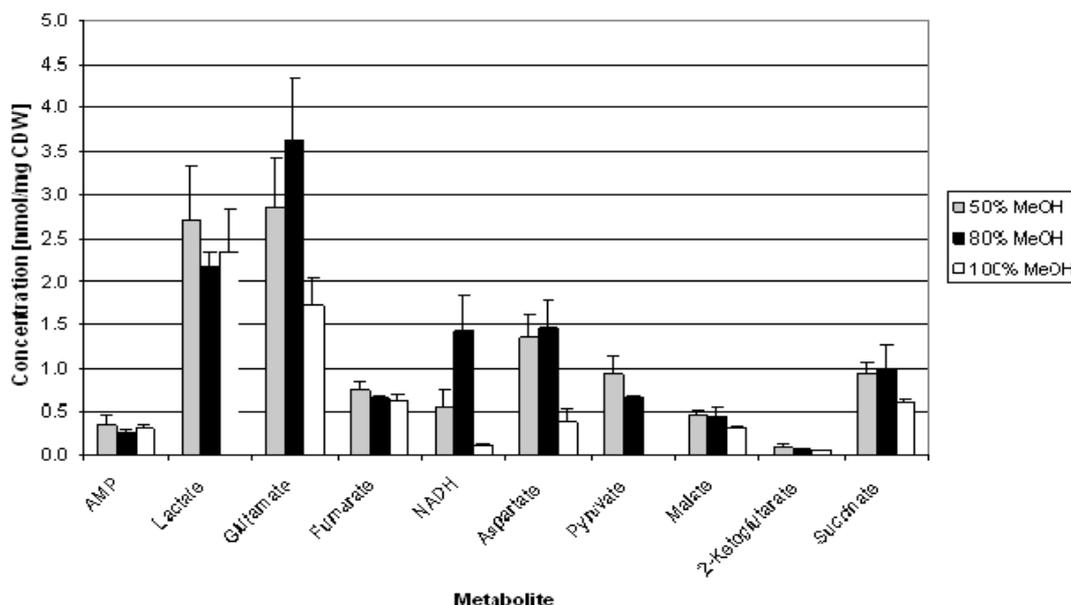
Figure 15: Effect of methanol:water ratio on extraction efficiency from E.coli cells. Cells were extracted in triplicates with different ratios of methanol:water (-20C) and analyzed with CE-MS.

7.3.2 Quantification in biological matrices

The validity of using calibration curves generated with aqueous metabolite standards for the quantitation of metabolites in biological samples was demonstrated by a matrix spike experiment. Briefly, metabolite standards at three different concentrations (15, 125 and 1000 μM) were added to the cold methanol extraction solution used for the extraction of *E. coli* cells. Each experiment was performed in triplicate. Linear regression analysis was performed and the calculated slopes were compared to those obtained with standards in pure water. The correlation between the slopes for the individual analytes is shown in Figure 16. A slope of 0.94 and a correlation coefficient (RSQ) of 0.97 proved that for most metabolites no significant matrix effects had occurred. Only ADP and to a lesser degree PG and F6P/G6P were slightly over- or underestimated by using the aqueous standard curves.
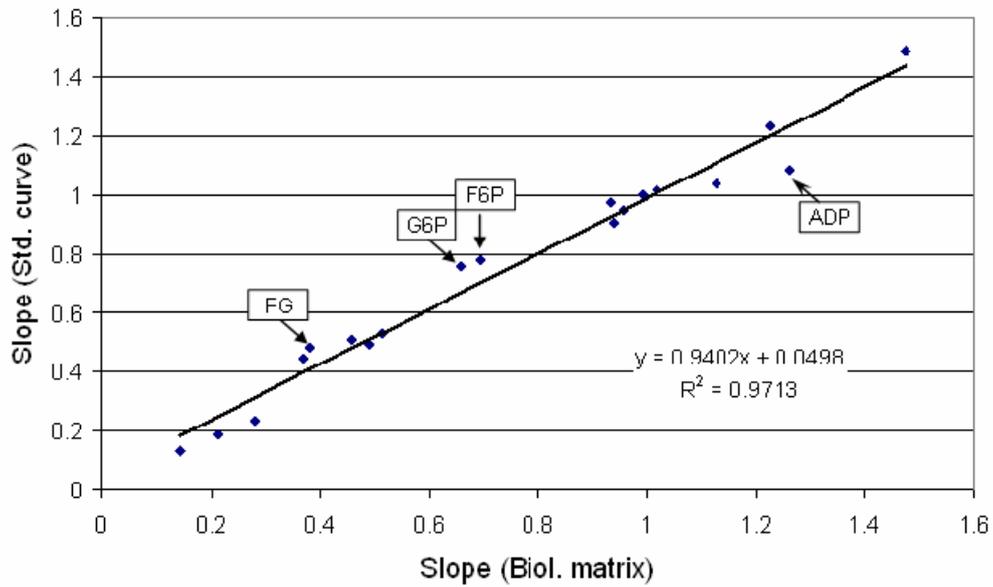
Figure 16: Comparison of the slopes of calibration curves generated by standards either prepared in water (y-axis) or in biological matrix (x-axis). *E.coli* cells were filtered and extracted in 80% cold methanol containing increasing amounts of the 20 metabolite standard mix (0 µM, 15 µM, 125 µM, 1000 µM).

8   Metabolome analysis of *E. coli*

8.1 Introduction

The metabolome of a microbe, while by far not as complex as that of a plant, still contains at least several hundred metabolites. For example the generation of *in silico* metabolomes, as described in chapter 4.3.8, of different microorganism species yielded 694 (*E. coli*), 537 (*B. subtilis*) and 458 (*S. cerevisiae*) different metabolites [96], i.e. low-molecular weight organic compounds with molecular weights of less than 1000 Da. For many of these metabolites, standards are not available commercially. It is therefore of great importance to develop methods that can screen for additional metabolites and provide high-confidence data leading to their identification.

In the previous two chapters, the development of a method for the robust quantification of negatively charged metabolites was described. To further evaluate the capabilities of the CE-TOF-MS method it was used to elucidate metabolic changes in the *Escherichia coli* deletion mutant *UdhA-PntAB* that lacks nicotinamide nucleotide transhydrogenase function [12], under both stationary and exponential growth conditions. The reproducibility of metabolite extraction and CE-TOF-MS analysis as well as the biological variance were evaluated.

In order to screen for additional metabolites, we adapted a freely available software tool, MZMine, for the specific challenges of CE-MS. We also established parameters to ensure the identification of potential metabolites with high probabilities. Finally, concomitant analyses with two different gas chromatography-mass spectrometry (GC-MS) methods allowed not only cross-validation of the quantitative results obtained by the various methods, but also led to a more comprehensive coverage of the *E. coli* metabolome. The results in this chapter, together with the findings from the previous two sections, were published in [83].

8.2 Material and Methods

8.2.1 Bacterial strains, growth conditions and sample extraction

The *E. coli* wild type strain MG1655 and the mutant UdhA-PntAB (MG1655 *ΔudhA ΔpntAB*) were cultured at 37ºC in 250 mL shake flasks containing 30 mL of Luria-Bertani (LB) or M9 minimal medium (0.1% NH_4Cl, 0.6% Na_2HPO_4, 0.3% KH_2PO_4, 0.05% NaCl) supplemented with 3 g/L glucose. After inoculation with a preculture in the same medium to an OD of 0.1, cells were harvested during strictly exponential growth or in stationary phase.

All other conditions were the same as described in chapter 7.2.

8.2.2 CE-TOF-MS

See chapter 6.2.1, optimized method parameters

8.2.3 Gas chromatography-mass spectrometry

An Agilent Technologies Model 6890 GC equipped with a 5975 Inert XL Mass Selective Detector (MSD) and an 7683B Auto Liquid Injector was used for the separation of trimethylsilylated metabolites on a RXI-5MS column (30 m x 0.25 mm ID x 0.25 µm film thickness, Restek, Bad Homburg, Germany). The initial oven temperature was set at 50°C, ramped at 8°C/min to 300°C, and held for 10 min. Helium was used as carrier gas at a flow-rate of 0.6 mL/min. The transfer line to the mass spectrometer was kept at 310°C. The mass spectrometer was operated in full scan mode from 50 - 600 m/z with a scan time of 0.5 s. Sample volumes of 1 µL were injected in splitless mode at 280°C.

For derivatization 50 µL of 10 mg/mL methoxylamine hydrochloride in pyridine were added and incubated at 60°C for 60 min, followed by 50 µL MSTFA for 60 min at 60°C.

Calibration was performed using the standard solutions described above. Hundred µL of the different calibration levels were transferred into a glass vial

with glass insert and 10 µL of the surrogate solution containing [U-^{13}C]lactate, [U-^{2}H]succinate, [U-^{13}C]glucose, cinnamic acid-d7 and norvaline (1mM each) was added. The standards were dried using a vacuum evaporator and derivatized as described above.

8.2.4 GC-MS of amino acids

An Agilent Technologies Model 6890 GC equipped with an 5975 Inert XL MSD, a PTV injector (Gerstel, Muehlheim, Germany) and a MPS-2 sample robot (Gerstel) was employed for amino acid analysis using a modified protocol from the Phenomenex EZ: faast kit. In contrast to the original protocol, the cation-exchange clean-up step was omitted. Amino acids were derivatized directly in the aqueous sample extract and extracted by isooctane. An aliquot of the organic extract (2.5 µL, 1:15 split) was injected into the GC-MS. Stable isotopes of amino acids were used as internal standards.

The derivatives were separated on a ZB-AAA column (Phenomenex Inc.), 15 m x 0.25 mm ID, 0.1 µm film thickness. The MS was operated in scan (mass range 50-420 m/z) and SIM (selected ion monitoring) mode. For quantification, calibration curves were generated using dilution series of standard solution (Phenomenex, Fluka). For most amino acids a calibration range of 0.75-675 µM was used.[127]

8.2.5 Data analysis for feature detection

For the quantitative analysis of selected metabolites, we used QuantAnalysis (Bruker Daltonics). Extracted ion chromatograms (EICs) were generated for the metabolites of interest, using their exact mass as given in the MetaCyc database [128] with a mass tolerance of ±0.01 Da. Calibration curves were measured before each batch of biological samples.

For the identification of unknown metabolites, we employed the 'Find All Compounds' algorithm provided with the DataAnalysis software suite from Bruker Daltonics. Metabolites found with this approach were identified by a database search using the exact mass ±0.01 Da in MetaCyc (http://metacyc.org/) and

the Human Metabolome Database (HMDB) (http://www.hmdb.ca/), and by comparison with the 'Generate Molecular Formula' (GMF) function, which estimates the potential molecular formula for a queried mass according to its exact m/z value and its isotopic pattern. The 'Find All Compounds' algorithm is able to look for correlations in the detected features, thus excluding isotopes and adducts. However, this can be problematic for the stable isotope labeled standards and for closely migrating compounds. Also, the software is not able to align the detected features of multiple samples. Therefore, we adapted an open-source data analysis software (MZMine, [109]) for the specific challenges of CE-MS analysis, *i.e.* we implemented the migration time correction method adapted from Reijenga et al. [124] as described in section 6.3.4. This allowed the rapid scanning of samples for additional metabolites. Alignment over multiple samples is possible; we included only those features in the alignment that were found in at least 6 out of 24 *E. coli* samples analyzed for every culture condition. The detected masses were queried as described above. The metabolites identified were validated by generating the corresponding EICs in DataAnalysis and applying the GMF tool to the molecular mass peak. Whenever possible, the identity of metabolites was confirmed with commercial standards.

8.3 Results and Discussion

8.3.1 Metabolic profiling in *E. coli* - method validation by comparison to other analytical methods

The optimized analytical protocol was applied to the comparative analysis of the *E. coli* wild type strain MG1655 with the double deletion mutant UdhA-PntAB (MG1655 *ΔudhA ΔpntAB*). The latter is deleted for both nicotinamide nucleotide transhydrogenases expressed in *E. coli*, which leads to a substantial re-routing of the metabolic flux from glycolysis to the pentose-phosphate pathway to meet the cellular demand for NADPH.[12] We were interested in the effect of this alteration in the main catabolic fluxes on steady-state cellular

metabolite levels. Initially, both strains were cultivated in complex medium and harvested after they had reached the stationary phase. Harvesting, metabolite extraction and sample preparation were performed as described in chapter 7. Four independent cultures were analyzed for each strain, with three replicates taken from each culture. The extracts were divided into two aliquots for CE-MS and GC-MS analysis, respectively. Of the 20 compounds in the quantitative metabolite panel (see Table 2), 18 were detected above the LLOQ in the extracts. Oxaloacetate and PG were not detected above the LLOQ. The analytical variability caused by extraction and CE-TOF-MS analysis, respectively, was assessed separately (see Table 3). The analytical variance was determined by measuring one of the *E.coli* samples four times over the course of one sample batch. It ranged from 3.7–12.6 % for all metabolites except for F6P and G6P, which did not resolve well in this biological matrix. In comparison, Soga et al. had reported an analytical variance of 0.7-9% for a metabolite standard mixture using CE-TOF-MS, and an inter-day variability averaging 30% for *E. coli* samples [97]. Extraction variance was assessed from three independent extractions of samples taken from one flask. It was between 8-17% for all metabolites except G6P, with 22.6%, which is similar to the analytical variance of this compound. Finally, the biological variability was determined by analyzing samples from four independent shake flasks for each of the two *E. coli* strains under investigation. Biological variability ranged from 3.4% to 31.3% and exceeded in most cases the experimental error.

To validate the CE-TOF-MS method, the *E. coli* samples were also subjected to two different GC-MS analyses: one was based on methoximation and silylation (GC-MS) for the analysis of small organic acids, while the other employed derivatization of primary and secondary amino groups with chloropropylformate for the analysis of amino acids (GC-(AA)-MS).[127] For GC-MS the analytical variance was smaller, between 0.9-4.6% for most of the metabolites, except for F6P, G6P and lactate (8-12%). This concords with other reports.[9] Extraction variability was similar to CE-TOF-MS (4.1-18.9%) as was the biological variance (3.6-30.3%). The analytical variance of GC-(AA)-MS (0.6-

11.3%) was similar to that of CE-TOF-MS, while extraction variability was between 9.9 – 20% for most amino acids except for glutamine (22.7%), proline, glycine and alanine (26.1- 29.4%), as well as tryptophan (44%). Biological variability was between 9.4-34% for most amino acids, exceptions were again alanine, glycine, proline, and tryptophan (35-70%). This suggested that extraction with 80% MeOH was not optimally suited for the latter amino acids, underscoring the difficulty of optimizing an analytical method for a great variety of metabolites. In conclusion, these findings demonstrated that the analytical variance of CE-MS and two different GC-MS methods was comparable, with GC-MS of methoximated and trimethylsilylated compounds performing slightly better than the other two methods.
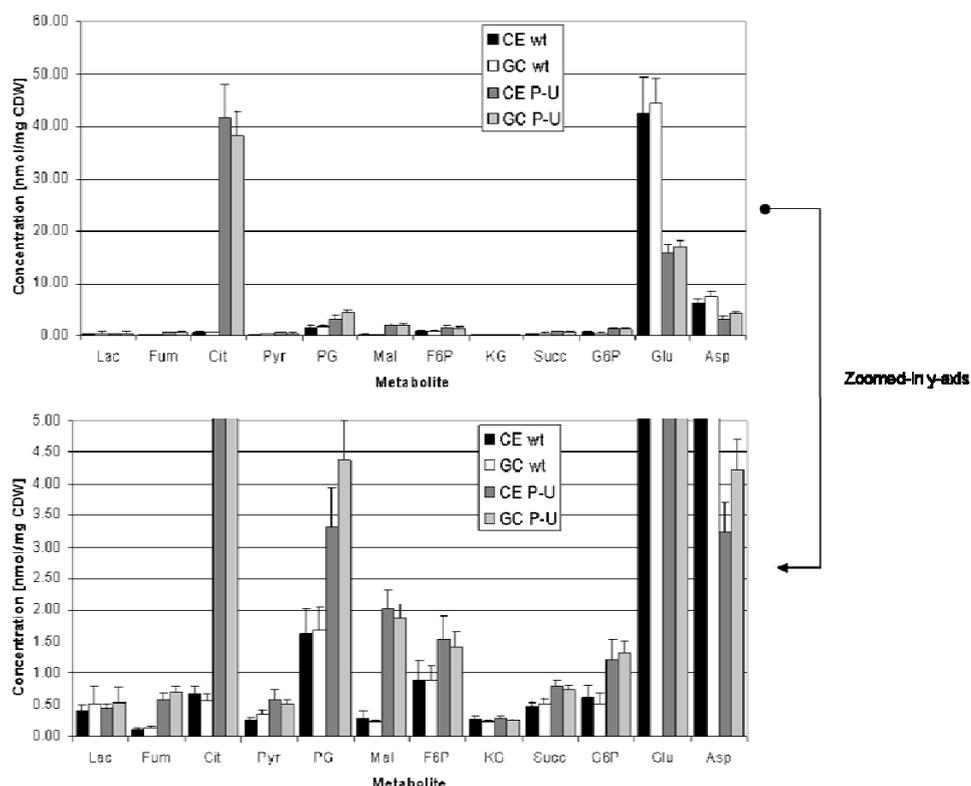
Figure 17: Comparison of the quantitative results of CE-MS vs. GC-(AA)-MS and GC-MS. Values are averages from 4 independent biological replicates, with 3 extracted samples per replicate.

For validation purposes, the quantitative data for metabolites detected in CE-MS and GC-MS was compared (Figure 17). Most metabolites showed an excellent concordance of the absolute quantities found in the cell extracts

(*e.g.*, lactate, pyruvate, malate, glutamate, citrate, fumarate, P-glycerate and succinate).

For AMP much higher amounts were found with GC-MS. The reason is that ATP and ADP are degraded to AMP under GC-MS conditions (Figure 18).
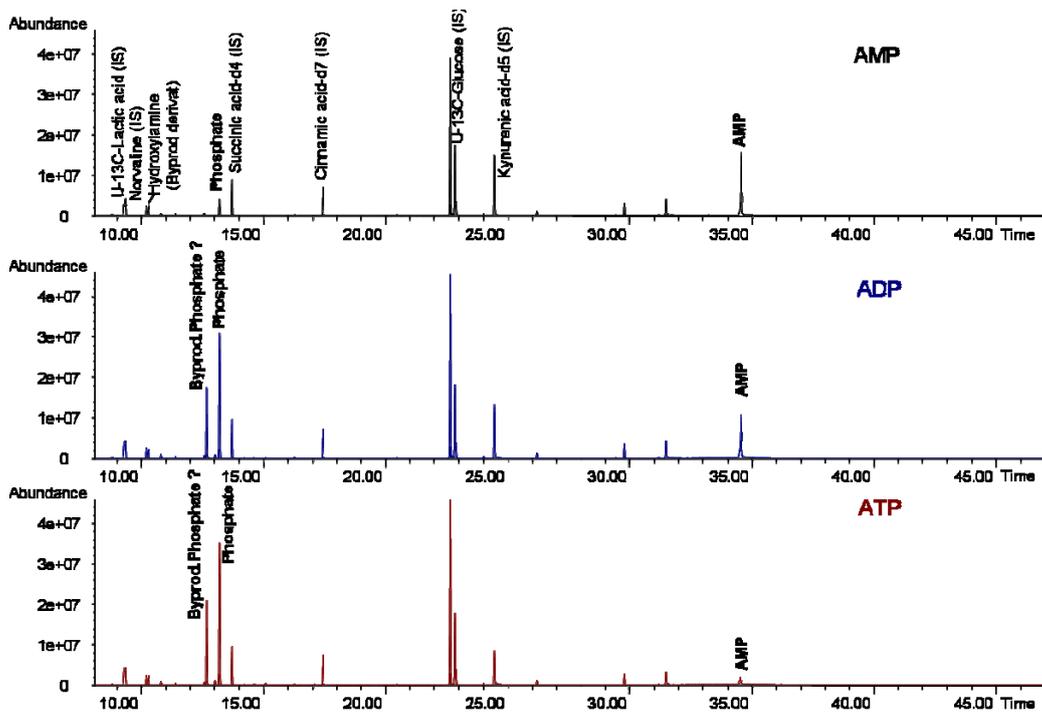
Figure 18: GC-MS chromatograms of AMP-, ADP- and ATP-standards. AMP cannot be quantified accurately, because ADP and ATP are converted to AMP during the analysis after loss of one or two phosphate groups.

For the amino acids glutamate and aspartate, which were detected by both CE-MS and GC-(AA)-MS, a very good correlation was observed (Figure 17). Of the other amino acids, many showed significant differences between the wild type and the mutant strain. For example, proline, contrarily to aspartate and glutamate, yielded six-times lower amounts in the wild type, while alanine, glycine and tryptophan were about two times lower. Aminoadipic acid was found at five times higher levels in the wild type.

All samples from one strain (n=12) were used to test for normal distribution by means of the Kolmogorov-Smirnov test of goodness of fit, where a p-value ≥0.05 indicates no deviation from normal distribution. All metabolites showed a normal distribution. We then applied the Student's t-test (two-tailed, pair wise)

to test for significant (P<0.05) differences in metabolite levels between the two strains. With the exception of lactate, 2KG, G1P and NAD, metabolite levels differed significantly between the wild type and the mutant strain (Table 4). Interestingly, the levels of most metabolites of the citric acid cycle were significantly higher in the mutant, while the amino acids glutamate and aspartate were higher in the wild type strain.

Table 4: Metabolite levels [nmol/mg cdw] in wild type *vs. PntAB-UdhA*, grown in complex medium and harvested in stationary phase.

Metabolite	Wild type (MG1655) [nmol/mg cdw] n=12	PntAB-UdhA [nmol/mg cdw] n=12	p-Value (t-test)
Lactate	0.41±0.08	0.44±0.11	0.12
Pyruvate	0.25±0.05	0.48±0.16	<0.001
Citrate	0.68±0.01	42.18±1.90	<0.001
2KG	0.28±0.05	0.30±0.03	0.2
Succinate	0.45±0.06	0.80±0.09	<0.001
Fumarate	0.11±0.03	0.58±0.11	<0.001
Malate	0.27±0.11	2.04±0.30	<0.001
Aspartate	7.92±1.13	4.24±0.31	<0.001
Glutamate	44.50±4.30	16.94±0.57	<0.001
G1P	0.09±0.009	0.10±0.01	0.057
G6P	0.62±0.20	1.20±0.19	<0.001
F6P	0.87±0.27	1.78±0.33	<0.001
FBP	0.14±0.03	0.75±0.35	0.012
3PG	1.62±0.36	3.37±0.61	<0.001
Gycerol-1P	0.36±0.00	0.55±0.09	0.0013
AMP	0.26±0.01	0.39±0.12	0.0015
ADP	0.75±0.14	0.96±0.28	0.014
NAD	4.63±1.15	4.25±0.44	0.22

8.3.2 Evaluation of the robustness of observed differences in metabolite levels

The results described above were obtained with *E. coli* cells that had been cultivated in complex medium. Thus, it was not clear whether the observed differences had resulted from altered flux distributions through the central carbon metabolism or whether they had been caused by interfering medium components. Therefore, we also grew both *E. coli* strains in minimal medium

supplemented with glucose. Samples were taken during exponential (log) and stationary (stat) phase to distinguish effects resulting from different growth phases. Four independent cultures were analyzed for each strain, with three replicates taken from each culture. The quantitative results are given in Table 5. Absolute metabolite levels differ as a function of growth stage, e.g. aspartate, 2KG resp. AMP. Further, cell culture medium related differences are found for FBP, which is only detected when cells are grown in the LB complex medium. Nevertheless, the absolute metabolite levels in *E. coli* are generally quite similar in cells cultured on different media and harvested at different growth stages. This is to be expected, because metabolite levels are influenced by a multitude of control mechanisms, which are able to balance internal and external perturbations.[97]

Table 5: Metabolite levels [nmol/mg cdw] in wild type *vs. PntAB-UdhA*, grown in minimal medium and harvested in both logarithmic and stationary phase.

| | M9 Logarithmic phase | | | M9 Stationary phase | | |
| | Wild type | PntAB-UdhA | | Wild type | PntAB-UdhA | |
Metabolite	Average content [nmol/mg cdw], n=12	Average content [nmol/mg cdw], n=12	p-Value (t-test)	Average content [nmol/mg cdw], n=12	Average content [nmol/mg cdw], n=12	p-Value (t-test)
Lactate	6.44±2.51	1±0.38	<0.001	0.96±0.16	0.83±0.15	0.047
Pyruvate	0.5±0.21	0.36±0.06	0.111	0.26±0.06	0.25±0.07	0.941
Citrate	1.03±0	16±3.93	<0.001	1.87±0.67	19.8±7.87	<0.001
2KG	< LLOQ	0.06±0.05	n/a	0.48±0.25	0.54±0.26	0.645
Succinate	0.75±0.32	1.12±0.41	0.022	0.92±0.21	1.8±0.49	<0.001
Fumarate	0.35±0.15	0.54±0.14	0.005	0.3±0.08	0.54±0.15	<0.001
Malate	0.29±0.21	1.02±0.2	<0.001	0.79±0.24	2.61±0.67	<0.001
Aspartate	2.07±0.62	2.34±0.34	0.242	6.09±1.57	3.78±1.27	0.002
Glutamate	23.1±5.57	12.3±2.81	0.002	68.9±22.6	32.4±9.1	0.004
G1P	< LLOQ	< LLOQ		0.18±0.04	0.15±0.04	0.133
G6P	0.87±0.32	2.5±0.69	<0.001	0.46±0.18	0.3±0.13	0.017
P-Glycerate	3.58±1.11	2.07±0.33	0.007	1.94±0.26	2.12±0.51	0.298
Gycerol-1P	3.23±3.8	0.51±0.22	0.027	1.84±1.26	0.68±0.37	0.016
AMP	0.84±0.19	0.59±0.09	<0.001	0.29±0.04	0.22±0.05	0.002
ADP	0.4±0.14	0.39±0.11	0.363	0.69±0.12	0.34±0.06	<0.001
NAD	2.04±0.79	3.03±0.71	0.004	4.26±0.73	3.77±1.08	0.22

In Table 6, the ratios of metabolite levels in wild type over the *PntAB-UdhA* strain under the different growth conditions are given. Glutamate levels were consistently higher (1.9-2.6) in the wild type under the three different conditions, while malate, succinate, fumarate and citrate (all part of the TCA) were much higher in the transhydrogenase deficient strain. Interestingly, aspartate was only elevated in wild type cells harvested during stationary phase, independent of the medium. NAD levels were constant in stationary phase, but reduced in the wild type strain during exponential growth. This probably mirrors an increased conversion to NADP, which is required for the biosynthetic reactions needed for cell division and growth.

Table 6: Relative changes in metabolite levels in wild type over *PntAB-UdhA* under different growth conditions.

Metabolite	LB (stationary)	M9 (exponential)	M9 (stationary)
Lactate	0.9	6.4*	1.2*
Pyruvate	0.5*	1.4	1
Citrate	0.016*	0.06*	0.09*
2-Ketoglutarate	0.9	n.d.[a)] in wt *	0.9
Succinate	0.6*	0.7*	0.5*
Fumarate	0.2*	0.7*	0.6*
Malate	0.1*	0.3*	0.3*
Aspartate	1.9*	0.9	1.6*
Glutamate	2.6*	1.9*	2.1*
G1P	0.9	n.d.[a)]	1.2
F6P	0.5*	0.3*	1.6*
FBP	0.2	n.d.[a]	n.d.[a)]
P-Glycerate	0.5*	1.7*	0.9
Gycerol-1P	0.7*	6.4*	2.7*
AMP	0.7*	1.4*	1.3*
ADP	0.8	1.0	2*
NAD	1.1	0.7*	1.1

* Significantly different between strains (p<0.05 in Student's t-test)
a) Below LLOQ

8.3.3 Metabolic fingerprinting of E. coli

The above results demonstrated the ability of our approach to yield validated quantitative data on a number of selected metabolites. However, the

challenges of modern systems biology often make it difficult to define in advance all of the metabolites that might be of interest in a certain biological context. To this end we made use of the TOF-MS's high mass accuracy and the fast scan speed over the complete mass range of 40-800 Da. The major problem in this regard is to extract meaningful conclusions from the huge amount of data collected in each electrophoretic run. Data extraction was performed using MZMine, an open-source program we adapted to the specific challenges of CE-MS, as described in section 2.8. With this approach more than 600 features, characterized by mass and MT and representing potential metabolites, were extracted from the electropherograms. A database search of the features with the exact mass ±0.01 Da in MetaCyc (http://metacyc.org/) and the Human Metabolome Database (HMDB) (http://www.hmdb.ca/) yielded 207 hits. After validation with the 'Generate Molecular Formula' tool integrated in the program Data Analysis from Bruker Daltonics, we identified 150 metabolites by their exact mass and their molecular formula, 71 of which could be confirmed by a reference standard (Table S1, Appendix). About 50 preliminarily identified metabolites had to be excluded, because no formula could be generated. Twenty-three were found to be either isotopes of abundant metabolites or multiply charged species. For most of the others, generating a reasonable formula was not possible due to interferences in the isotope peaks, either by overlapping compounds or a high background. These findings demonstrated the importance of integrating the identification of metabolites via their exact mass with additional quality control parameters. These included their molecular formula, corrected migration time and, if possible, comparison with a commercially available standard. Using the criteria of exact mass, molecular formula and migration time, we have as yet not assigned incorrectly a metabolite that could not be confirmed by measuring the corresponding standard substance.

Peak detection and peak area integration in MZMine are done automatically. Although integration markers cannot be altered manually, we found a very good correlation for the wt:*PntAB-UdhA* ratios between the metabolites

quantified with QuantAnalysis and the MZMine derived data for most analytes. Exceptions were the hexose-phosphates, which are not resolved in MZMine, and pyruvate and 2KG in M9 log, which are present at low concentrations. Less stringent constraints for automated peak detection might give better results for low-abundance metabolites. In turn, however, this may lead to an increased number of false-positive features. We also did a semi-quantitative analysis in terms of peak area relative to internal standard for 12 randomly chosen additional metabolites that were identified in the extract and again the ratios were highly similar to the ones obtained with MZMine (Table 7).

Table 7: Comparison of the metabolite ratios in wild type over *PntAB-UdhA* obtained by the signal integration programs QuantAnalysis and MZMine.

Metabolite	LB (Stationary)		M9 (Exponential)		M9 (Stationary)	
	Quant Analysis	MZMine	Quant Analysis	MZMine	Quant Analysis	MZMine
Lactate	0.9	1.3	6.4*	4.07*	1.2*	1.3*
Pyruvate	0.5*	0.3*	1.4	n.d.[a)]	1	0.3
2KG	0.9	1.0	n.d.[a)] in wt	0.6	0.9	1.1
Succinate	0.6*	0.7*	0.7*	0.7*	0.5*	0.5*
Glutamate	2.6*	2.8*	1.9*	2.1*	2.1*	2.3*
G1P/F6P[b)]	0.9/0.5*	0.5*	n.d.[a)]/0.3*	0.4*	1.2/1.6*	1.4*
ADP	0.8	0.8	1.0	1.0	2*	2.0*
NAD	1.1	0.8	0.7*	0.7*	1.1	1.1
Benzoic acid	2.4*	2.7*	2.1*	2.2*	1.1	1
cis-Aconitic acid	0.03*	0.01*	PU only*	0.1*	0.3*	0.2*
Glucuronate	0.3*	0.04*	PU only*	0.1*	PU only*	0.1*
Glutamine	1.6	n.d.[a)]	0.7*	0.3*	1.6*	1.7*
Glutarate	1.2	1.2	0.7*	0.6*	2.4*	2.2*
Glutathione	1.9*	1.9*	1.6*	1.6	1.2*	1.4*
NADP	1.8*	1.5	1.1	1.6	3.2*	5*
Orotic acid	PU only*	0.3*	wt only*	11.6*	wt only*	n.d.[a)]
Ribose-5-P	0.7*	0.5	2.2	n.d.[a)]	0.9	0.7
Sedoheptulose-7-P	0.6*	0.6*	1.1	1	1.1	1.1
Tartaric acid	1.1	0.6	1.2	n.d.[a)]	n.d.[a)]	n.d.[a)]
UDP-D-glucose	1.1	0.8*	1.3*	1.2	2.6*	2.4*

* Significantly different between strains (p<0.05 in Student's t-test)
a) Below LLOQ
b) G1P, F6P and G6P are not separated in the automated MZMine analysis

Although the electrophoretic conditions had been optimized for the separation of negatively charged metabolites, it proved possible to detect even positively charged compounds, as they were swept to the capillary outlet by the EOF and the pressure applied after the electrophoretic separation for washing and conditioning. Despite their incomplete resolution, they could be in part quantified accurately as exemplified for such metabolites as valine, lysine, proline, nicotinamide and glucose (in comparison with GC-MS derived data, data not shown).

In summary, we found higher levels of metabolites of the pentose-phosphate pathway (PPP) and the citric acid cycle in the transhydrogenase-deficient mutant strain. The differences in PPP were only seen in LB medium. In minimal medium, the PPP intermediates were detected at low levels and, therefore, the resulting higher standard deviations might obscure any significant differences. The increased abundance of PPP intermediates in the deletion mutant might indicate an increased flux of substrate through these pathways to fulfill the cell's need for NADPH, which is a cofactor in many reductive synthetic reactions such as fatty acid and nucleic acid synthesis. NADPH is produced in the PPP by the enzyme 6-phospho-D-gluconate:$NADP^+$ 2-oxidoreductase and in the citrate cycle by the enzyme isocitrate:$NADP^+$ oxidoreductase. In addition, we consistently found increased levels of reduced glutathione (GSH) in the wild type strain under all conditions. This might be due to the decreased availability of NADPH, which is necessary for the reduction of glutathione disulfide (GSSG) to GSH by the enzyme glutathione reductase. Along the same line the differences in glutamate levels could be explained. Glutamate is synthesized from 2KG by the enzyme glutamate synthase, which is NADPH dependent. Lower levels of NADPH could thus lead to lower reaction rates. A summary of the data is given in Figure 19, showing significant differences in the detected metabolite levels (t-test, $P<0.05$) integrated into a central carbon metabolism pathway map.
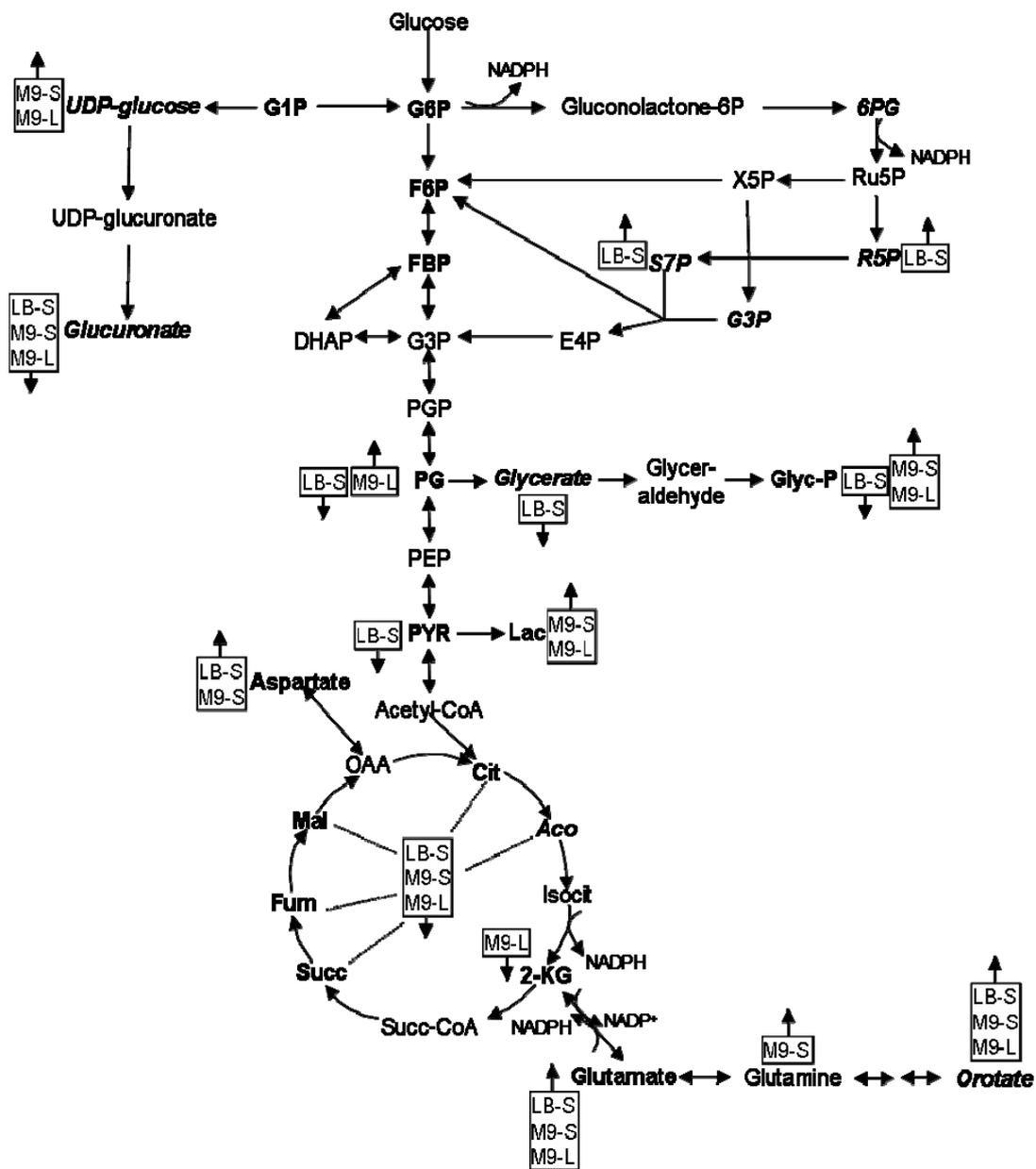
Figure 19: Diagram of the central carbon metabolism. Metabolites that were quantified with calibration curves are marked in **bold**, while metabolites measured semiquantitatively are marked in **bold** and *italic*. Metabolites that are found in higher concentrations in the wild type are indicated with an arrow pointing up; higher concentrations in the *PntAB-UdhA* strain are indicated with an arrow pointing down; the growth conditions are abbreviated as LB-S, complex medium stationary phase; M9-S, minimal medium stationary phase; and M9-L, minimal medium logarithmic phase.

71

8.3.4 The validated CE-TOF-MS method in the context of recent developments in metabolome analysis

We compared our *E. coli* metabolite data with a CE-TOF-MS method for the analysis of negatively charged metabolites using commercial coated SMILE capillaries and positively charged metabolites as well as nucleotides, separated in untreated fused silica capillaries employing different BGE and preconditioning.[97] In total, Ishii et al. observed 61 of the 150 high-confidence identifications reported in our study (40.7%). If only negatively charged metabolites were considered, we detected 29 of 62 reported by Ishii et al. (46.8%). However, these numbers increased to a concordance of about 70% (26/38) upon comparison of metabolites that had been found in at least 20% of replicate samples. The differences in metabolites observed in the two studies might be due to the slightly different genetic backgrounds of the wild type strains used, namely MG 1655 vs. BW 25113, and the different culture conditions employed (steady-state bioreactor vs. shake flask).

In comparison to other methods developed for the quantitative analysis of the intermediates of the central carbon metabolism, CE carries the advantages that only minute sample amounts are required and that there is no need for pre-column derivatization of analytes. A disadvantage is the variability of migration times, but there are, as reported here, effective ways of correcting for such variation. The GC-MS methods we used in this project yielded LODs comparable to those of CE-TOF-MS (data not shown). LC-MS/MS based methods achieve lower LODs, typically in the concentration range of 1-50 nM [11] as is also shown in the next chapters. However, one has to keep in mind that for a TQ-MS to yield the highest possible sensitivity, the instrument has to be operated in multiple-reaction monitoring (MRM) mode, which requires prior knowledge of the analytes. Screening for additional metabolites would therefore require an additional analytical run, preferably using a TOF-MS, since quadrupole instruments are less sensitive in scan mode and give only unit mass resolution, which leads to a high number of possible metabolite hits.

9 Method development for rapid lactate measurement

9.1 Introduction

The generation of lactate from pyruvate is one possibility for cells to regenerate NAD, which is consumed during glycolysis when glyceraldehyde-3-phosphate is converted to 1,3-bisphosphoglycerate. Lactate is secreted to the cell medium and accumulates there. Alterations in lactate levels can yield information on alterations in the central carbon metabolism, for example the production of lactate is associated with the rapid growth of tumor cells.[129] In several projects we are especially interested in the excretion or uptake of lactate and its isotopes. For these projects we wanted to develop a method that is able to quantify lactate and all possible isotopes of this compound. The main focus during method development was on analysis speed, because we needed a method that was able to screen a high number of samples in a reasonable amount of time.

Method development was based on a protocol described by Luo et al. [11] for IP-LC-MS/MS. This method is based on the formation of ion pairs between the negatively charged analytes and the buffer additive tributylamine (TBA), which allows their separation on a polar endcapped reversed phase column. Very good resolution of a number of charged compounds was reported, however the complete method takes 90 minutes for the analysis of one sample. We could shorten the method to a total analysis time of 16 minutes, while still retaining the separation of lactic acid and several other metabolites of interest, such as pyruvate and the hexose phosphates. The quantitative capabilities of this approach are demonstrated.

The analysis of lactate uptake into cells of the immune system was published in [114].

9.2 Material and Methods

9.2.1 IP-LC-MS/MS

An Agilent 1200 SL HPLC system was used for liquid chromatography. The separation was performed on a Synergi Hydro-RP column (150 mm x 2.0 mm I.D., 4 µm particles, 80 Å pore size). Eluent A was 10 mM tributylamine (TBA), 15 mM acetic acid (pH 4.95), eluent B was pure methanol. The optimal flow rate was found to be 350 µl/min, with the column held at a constant temperature of 50ºC. The optimized gradient conditions were as follows: 0 – 7 min from 0% to 90% B and hold for 3 min at 90% B. Before every run the column was equilibrated at 100% A for 6 minutes.

The HPLC system was directly coupled to a 4000 QTRAP mass spectrometer (ABI/MDS Sciex, Concord, Canada). The MS was operated in negative ion mode with selected reaction monitoring (SRM). Source settings were as follows: IonSpray Voltage (IS) -3700 V, auxiliary gas temperature (TEM) 350ºC; curtain gas (CUR) nebulizer gas (GS1), auxiliary gas (GS2) were set to 10, 50 and 50 (arbitrary units), respectively; collision gas (CAD) was set to medium. The entrance potential (EP) was set to -10 V for each compound. Declustering potential (DP), collision energy (CE) and collision cell exit potential (CXP) were optimized for lactate and pyruvate, as well as the MS/MS fragment pattern (see Table 9 in the next chapter). Compound specific optimization was performed by direct infusion of a 10 µM solution of each standard substance in water/methanol 50:50 (v/v), containing 10mM TBA and 15mM acetic acid. The infusion was performed at a flow rate of 10 µl/min using a model 11 PLUS syringe pump (Harvard Apparatus, Holliston, USA).

9.3 Results and Discussion

9.3.1 Method development for fast lactate measurement

We modified the IP-LC-MS/MS method described by Luo et al. [11] for the rapid analysis of lactate and pyruvate in biological samples. Using the original gradient, lactate, pyruvate and several other metabolites are eluted from the column under isocratic conditions at 100% A. However, the isocratic part of the method including washing and equilibration takes about 30 minutes. In a first step, we increased the flow rate from 200 µL/min to 350 µL/min. A sample chromatogram of the isocratic run at 350 µL/min of a mixture of metabolite standards is shown in Figure 20 a. In order to further decrease total analysis time, different gradients and their effect on the chromatographic behavior of lactate were tested (Table 8).

Table 8: Effect of different gradient conditions on the chromatographic behavior of lactate

Gradient	Peak Area [x10^6]	FWHM	RT [min]
100% A, 15 min	0.47	0.28	8.9
95% A, 15 min	0.76	0.26	6.7
0 --> 25% B, 15 min	1.02	0.19	7.3
0 --> 50% B, 15 min	1.48	0.17	6.6
0 --> 60% B, 7 min	1.91	0.15	6.0
0 --> 90% B, 11 min	1.95	0.14	6.1
0 --> 90% B, 7 min	2.19	0.14	5.7

The optimal gradient was 0 → 90% B in 7 minutes, thus reducing the run time by half to an injection every 16 minutes. A chromatogram under optimized gradient conditions is shown in Figure 20 b. The benefits of the fast gradient as compared to the isocratic separation for lactate were as follows: retention decreased from 8.9 min to 5.7 min; FWHM was halved from 0.28 min to 0.14 min; while the area increased by a factor of about 4, from 2.73E+05 to 1.23E+06.
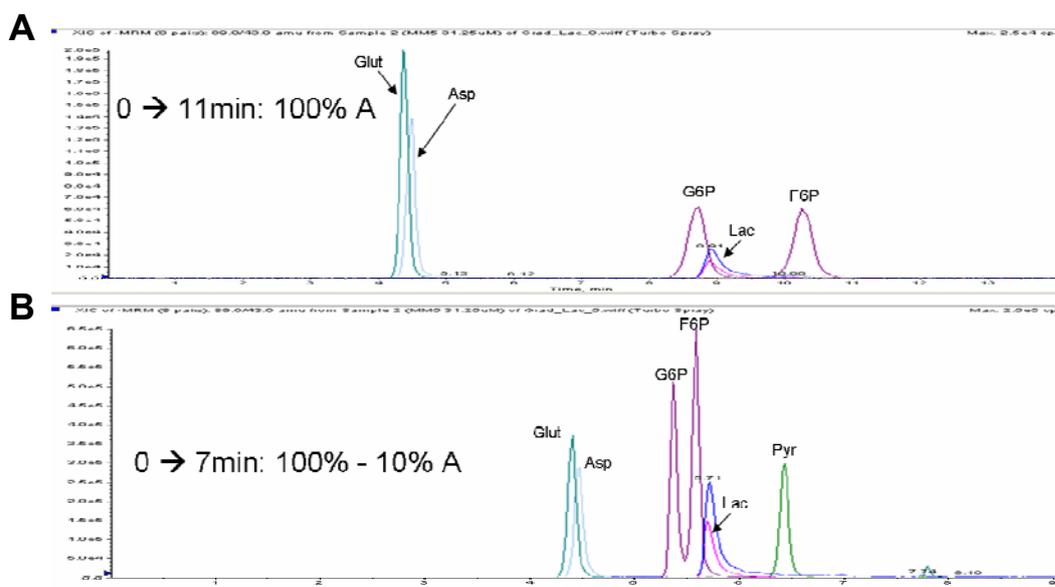
Figure 20: Effect of gradient optimization on the separation of selected compounds.

9.3.2 Uptake of lactic acid into different immune cell populations

The modified IP-LC-MS/MS method was applied to the measurement of lactate uptake into different immune cell populations, such as cytotoxic T-lymphocytes (CTLs), T-helper cells, regulatory T-cells and monocytes. The lactate isotopomer [3-^{13}C]lactate was added in varying concentrations to the different immune cell cultures. Cells were extracted and the concentration of labeled lactate in the cell extracts was determined. By this approach, uptake of lactic acid into CTLs and its increase upon extracellular acidification was demonstrated (Figure 21).[114]
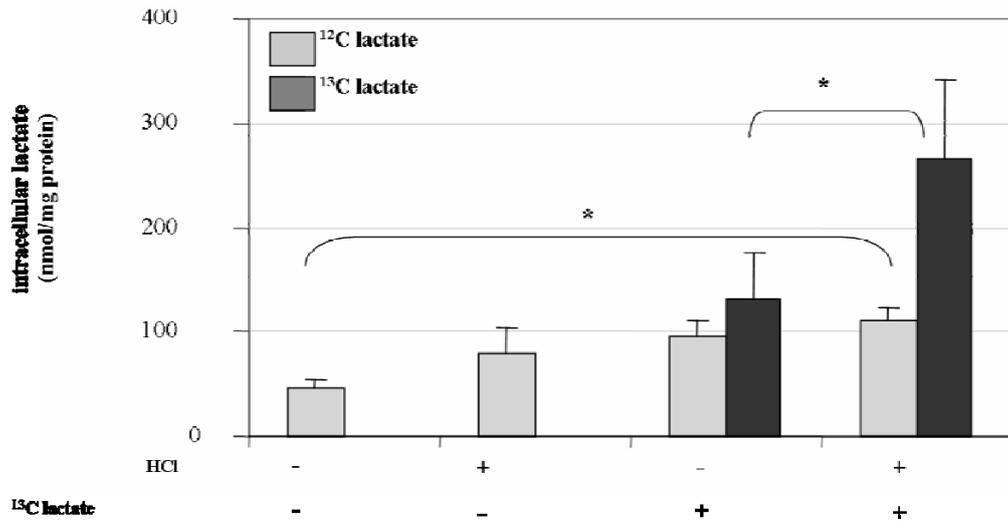
Figure 21: Uptake of ^{13}C-labelled lactate by CTL. CTL were incubated for 30 minutes with 20mM external ^{13}C-lactate in the presence or absence of HCl (n=4). Endogenous lactate and the uptake of exogenous lactate was determined in the cell lysates by CE-TOF-MS.[114]

Furthermore, the method was used to study the effects of lactic acid on mono-cytes. It was shown that lactic acid inhibits TNF-alpha secretion by suppressing glucose metabolism in monocytes.[130]

10 Method development for targeted metabolite analysis using rapid resolution IP-LC-MS/MS

10.1 Introduction

The method described by Luo et al. [11], while achieving a high resolution of critical isomers, e.g., G6P and F6P, takes 90 minutes. This is not suitable for the high-throughput analysis of the dozens to hundreds of biological samples typically required for a sound statistical analysis. The goal in this subproject was to substantially accelerate the analysis, while still retaining acceptable resolution of critical isomers. This was implemented in several steps: (i) use of smaller particles (2 µm) to allow shorter columns and increased flow rate; (ii) optimization of gradient conditions with regard to short retention time, resolution, peak shape and sensitivity; (iii) robust quantification of selected metabolites by introduction of stable isotopes.

The Van Deemter curve relates the height of a theoretical plate (H) to the linear velocity of the mobile phase (u). The smaller H the better the separation, because more peaks can be resolved on a column (N=L/H). The resolution is directly proportional to the square root of N, thus the smaller H the better the resolution. The factors that mainly influence H are the size of the stationary-phase particles and the diffusion coefficient. From the Van Deemter plot (see Figure 7) it becomes clear, that the smallest values for H are reached at a certain flow rate optimum. It was shown, that this optimum is substantially dependent on the particle size; it is shifted to larger flow rates with decreasing particle size. Also, the optimum flow region becomes broader under these conditions, allowing the use of higher flow rates.

We evaluated the benefits of the rapid resolution approach for the IP-LC-MS/MS analysis of intermediates of the central carbon metabolism and established the quantitative capabilities of the optimized method. The method was further validated for biological applications by measuring *E. coli* cell extracts.

10.2 Material and Methods

10.2.1   LC-MS/MS

An Agilent 1200 SL HPLC system was used for liquid chromatography. The separation was performed on a YMC-Ultra HT Hydrosphere C18 column (100 mm x 2.0 mm I.D., 2 μm particles, 120 Å pore size). Eluent A was 10 mM tributylamine (TBA), 15 mM acetic acid (pH 4.95), eluent B was pure methanol. The optimal flow rate was found to be 380 μl/min, with the column held at a constant temperature of 40ºC. The optimized gradient conditions were as follows: 0 – 5 min hold at 100% A, 5 – 25 min from 0% to 90% B and hold for 5 min at 90% B. Before every run the column was equilibrated at 100% A for 7 minutes.

The HPLC system was directly coupled to a 4000 QTRAP mass spectrometer. The MS was operated in negative ion mode with selected reaction monitoring (SRM). Source settings were as follows: IonSpray Voltage (IS) -3700 V, auxiliary gas temperature (TEM) 350ºC; curtain gas (CUR) nebulizer gas (GS1), auxiliary gas (GS2) were set to 10, 50 and 50 (arbitrary units), respectively; collision gas (CAD) was set to medium. The entrance potential (EP) was set to -10 V for each compound. Declustering potential (DP), collision energy (CE) and collision cell exit potential (CXP) were optimized for each compound, as well as the MS/MS fragment pattern (Table 9). Source settings and compound specific optimization were performed by direct infusion of a 10 μM solution of each standard substance in water/methanol 50:50 (v/v), containing 10mM TBA and 15mM acetic acid. The infusion was performed at a flow rate of 10 μL/min using a model 11 PLUS syringe pump. Additional SRM transitions were analyzed for all possible isotope distributions of selected compounds. Care was taken to keep the scan time for each compound at or above 30 ms. In order to still record enough data points for each peak (cycle time < 1.5 s), the run was subdivided into 4 periods, which were selected according to retention times of the compounds of interested.

Table 9: Optimized SRM parameters for measured metabolites

Metabolite	Mass [Q1]	Mass [Q3]	Main product ion	DP [V]	CE [V]	CXP [V]
AMP	346	79	$[PO_3]^-$	-88	-57	-11
Succinate	117	73	$-CO_2$	-43	-16	-10
Malate	133	115	$-H_2O$	-20	-15	-18
Lactate	89	43	-HCOOH	-47	-18	-5
Pyruvate	87	43	$-CO_2$	-42	-13	-4
P-glycerate	185	97	$[H_2PO_4]^-$	-45	-21	-14
Aspartate	132	88	$-CO_2$	-25	-17	-13
FBP	339	97	$[H_2PO_4]^-$	-58	-29	-10
G6P	259	79	$[PO_3]^-$	-62	-57	-22
F6P	259	79	$[PO_3]^-$	-57	-62	-11
Glutamate	146	102	$-CO_2$	-51	-19	-16
NAD	662	540	-Nicotinamide	-60	-19	-12
NADP	742	620	-Nicotinamide	-53	-22	-13
Fumarate	115	71	$-CO_2$	-45	-11	-10
Glycerol-P	171	79	$[PO_3]^-$	-55	-21	-11
2-Ketoglutarate	145	101	$-CO_2$	-35	-12	-16
ADP	426	79	$[PO_3]^-$	-80	-88	-3
cis-Aconitate	173	85	$-2CO_2$	-30	-17	-12
Ribose-5P	229	97	$[H_2PO_4]^-$	-52	-18	-14
DHAP	169	97	$[H_2PO_4]^-$	-32	-14	-11
Lactoylglutathione	378	306	Glutathione	-78	-22	-15
Methylglyoxal	71	43	-CO	-80	-13	-4
Glyoxylate	73	45	-CO	-65	-13	-4
Acetyl-CoA	808	79	$[PO_3]^-$	-115	-120	-11

10.2.2 Samples and sample preparation

Cultivation on minimal medium and extraction of *E. coli* cells was performed as described above. The internal standard solution was the same as described below in the quantification section.

10.2.3   Quantification

Absolute quantitation of 21 different metabolites, including 2KG, F6P, G6P, FBP, AMP, ADP, NAD, NADP, glutamate, aspartate, glycerol-1P, glycerol-2P, fumarate, cis aconitate, lactate, pyruvate, malate, succinate, citrate, 3PG and ribose-5P was accomplished with a dilution series of an equimolar (5 mM) aqueous stock solution of all standards (= standard mix), to which 10 resp 2 µM each of the extraction and internal standard compounds [2,3,3,3-$^2H_4$]lactate, [U-^{2}H]succinate, [2,3,3-$^2H_3$]malate, [2,3,3,4,4-$^2H_5$]glutamate, [U-^{13}C]fumarate, [U-^{13}C]pyruvate, [U-^{13}C]glucose-1-phosphate and [U-^{13}C, U-^{15}N]AMP had been added. Prior to analysis of biological samples, 15 calibration points were generated over a concentration range of 2.5 nM -100 µM. Spectral peak integration was done with Analyst 1.4.2 (Applied Biosystems) following normalization of the peak areas of the standards to the area of the closest migrating internal standard. Concentrations were then inferred from the calibration curves. Stability of measurement was checked by periodical injection of a quality control sample between biological samples.

10.3 Results and Discussion

10.3.1   Optimization of flow rate – Van Deemter Plots

Van Deemter plots are generated under isocratic conditions. We, therefore, used 100% eluent A to elute several standard compounds (aspartate, glutamate, lactate, G6P, F6P and pyruvate) that can be separated with this solvent strength. The flow rate was ramped from 100 µL/min to 400 µL/min in increments of 50 µL/min. It would have been interesting to test lower flow rates, but this was not possible in a stable fashion with the LC pump used. At higher flow rates the pressure was quite close to the upper pressure limit of the column. Figure 22 depicts the experimentally determined Van Deemter plots for selected compounds.
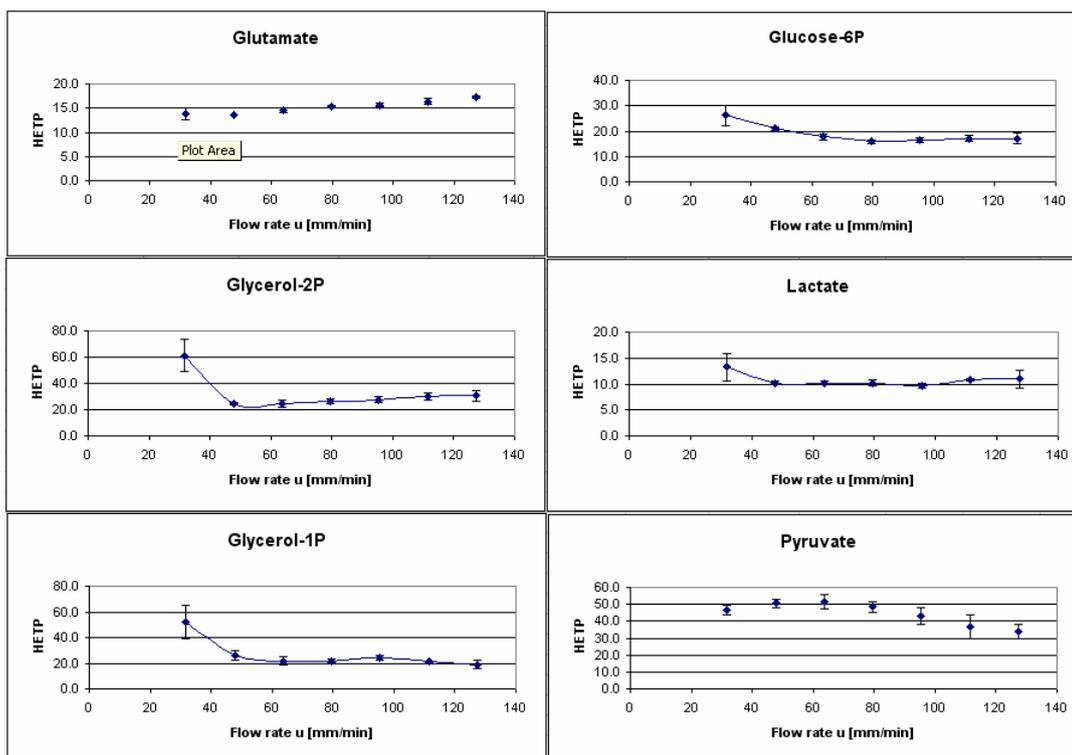
Figure 22: Van Deemter plots for different compounds that are eluted under isocratic conditions at 100% A.

It is interesting, that the effect of the flow rate on H differs for compounds with different k-values. For glutamate, H does not change at higher flow rates. The reason for this might be that glutamate is very weakly retained under the applied conditions and, thus, is not much affected by changes in flow rate. However, the increase in flow rate results in a proportional decrease in retention time and peak width, i.e. a fourfold increase in flow rate from 100 to 400 µL/min results in a roughly fourfold reduction in retention time and peak width (Table 10).

Table 10: Effect of flow rate changes on retention time and peak width.

Flow rate u [µl/min]	Glutamate		Glucose-6P		Lactate	
	RT [min]	FWHM [min]	RT [min]	FWHM [min]	RT [min]	FWHM [min]
100	8.0	0.22	12.3	0.46	13.0	0.34
150	5.4	0.15	8.2	0.28	8.7	0.21
400	2.2	0.07	3.3	0.10	3.5	0.09

The increase in efficiency at higher flow rates allows for the improved resolution of isobaric intermediates. These are glucose-6-phosphate (G6P), fructose-6-phosphate (F6P) and glucose-1-phosphate (G1P) respectively glycerol-1P and glycerol-2P. As can be seen from Figure 23, the resolution increases for flow rates up to 200 µL/min and remains at that level for higher flow rates, even though retention times decrease by 50%.
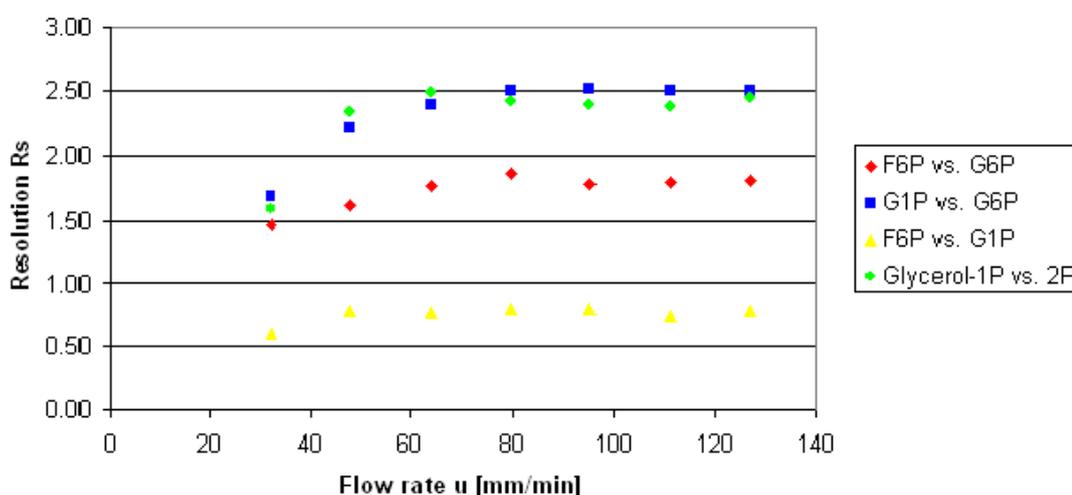
Figure 23: Effect of flow rate on the resolution of critical pairs of isobaric glucose metabolites.

G1P was not used in the standard mix during further analysis, because the hexose-phosphates have very similar calibration parameters and G1P was not found previously in *E. coli* samples grown on minimal medium and harvested during exponential growth as described in chapter 8.3.2 and [83].

10.3.2   Optimization of gradient conditions

A number of different gradient conditions were tested (Table 11) in order to find an optimized method that allows for the fast and sensitive detection of the metabolites of interest. It was clearly seen that speeding up the gradient led to narrower peaks, which increased peak area and, thus, sensitivity. Nevertheless, we kept the initial isocratic plateau at 100% A, albeit for a shortened period of time, in order to maintain the high resolution of isobaric compounds such as the hexose- and pentose-phosphates as well as the glycerol-

phosphates. As already seen for lactate and pyruvate, a steeper gradient had a positive effect on FWHM and thus signal intensities for most metabolites (data not shown). Gradient RR9 was finally chosen for subsequent experiments after a final optimization step for several ion source parameters (CUR, TEM). The optimized method amounted to a total run time of 37 minutes, which results in a reduction of total analysis time by 59% as compared to the method initially described by Luo et al.

Table 11: Gradient conditions used for gradient optimization in rapid resolution IP-LC-MS/MS. Equilibration was performed for 10 min before each run with the initial solvent conditions.

Gradient name	Gradient conditions
RR 1	0-15': 100→80% A; -35': 80→65% A; -45': 65→40% A; -50': 40→10% A; -55': 10% A
RR 2	0-5': 100→80% A; -35': 100→10% A; -40': 10% A
RR 3	0-5': 100% A; -10': 100→85% A; -20': 85→70% A; -35': 70→10% A; -40': 10% A
RR 4	0-5': 100% A; -10': 100→85% A; -15': 85→75% A; -35': 75→10% A; -40': 10% A
RR 5	0-5': 100% A; -10': 100→85% A; -20': 85→70% A; -30': 70→10% A; -35': 10% A
RR 6	0-5': 100% A; -10': 100→85% A; -15': 85→75% A; -25': 75→10% A; -30': 10% A
RR 7	0-5': 90% A; -20': 90→10% A; -25': 10% A
RR 8	0-5': 100% A; -10': 100→85% A; -15': 85→70% A; -25': 70→10% A; -30': 10% A
RR 9	0-5': 100% A; -25': 100→10% A; -30': 10% A

This substantial reduction in total analysis time represents a very important prerequisite for the robust analysis of the large batches of biological samples that typically have to be investigated in system biological studies.

10.3.3   Evaluation of quantitative capabilities

For the absolute quantification of selected metabolites of interest, calibration curves were generated. To that end, an equimolar mixture of metabolites of interest was diluted serially over a concentration range of 100 µM to 0.1 nM. Calibration curve parameters as well as LOD and limits of quantitation are presented in Table 12.

Table 12: Calibration curve parameters for the optimized RR-IP-LC-MS/MS method.

Metabolite	RSQ	LOD [nM]	LLOQ [nM]	ULOQ [µM]	Variance, intra-day (%) 25 µM / 500 nM
2KG	0.997	50	250	100	5.5 / 8.6
ADP	0.9946	10	250	100	2.2 / 1.5
AMP	0.9995	25	50	100	1.7 / 1.4
Aspartate	0.9987	0.5	1	100	0.7 / 0.3
cis Aconitate	0.9966	25	50	100	1.4 / 2.5
F6P	0.9971	1	2.5	50	0.7 / 0.5
FBP	0.9416	5	100	100	2.9 / 2.0
Fumarate	0.9985	10	50	100	1.1 / 1.7
G6P	0.9978	2.5	5	50	1.4 / 1.2
Glutamate	0.9959	1	2.5	50	0.5 / 0.3
Glycerol-1P	0.9976	2.5	5	100	0.5 / 1.3
Glycerol-2P	0.9987	2.5	5	100	0.9 / 2.7
Lactate	0.9969	50	100	100	2.1 / 4.0
Malate	0.9967	10	25	100	0.5 / 1.6
NAD	0.9953	0.2	2.5	50	1.4 / 1.1
NADP	0.997	5	25	100	0.9 / 1.3
P-Glycerate	0.9987	5	50	100	1.5 / 0.7
Pyruvate	0.9995	250	500	100	1.1 / 6.7
Rib-5P	0.9981	1	2.5	50	0.6 / 0.9
Succinate	0.9964	25	50	100	1.0 / 1.4

The LOD was defined as the concentration, at which a compound can be detected with a signal-to-noise (S/N) of ≥3. The LLOQ and ULOQ were defined according to the FDA Guide for Bioanalytical Method Validation [126] as the lowest and highest points of the calibration curve, respectively, at which a compound is measured with an accuracy between 80-120%. Intra-day reproducibilities were evaluated by consecutive injections of standards at two different concentrations. Average RSDs (n=5) were 1.4% (0.5–5.5%) for the 25 µM standard and 2.5% (0.3-11.3%) for the 500 nM standard. Inter-day reproducibilities were evaluated by measuring standards and biological samples several times over the course of several sample batches for approximately three days. Average RSDs (n=5) were 10.8% for the *E. coli* sample, 4.2% for the 25 µM standard and 9.2% for the 500 nM standard.

Evaluation of the inter-day reproducibility shows the beneficial effect of using isotope-labeled internal standards. Upon comparison of the average repro-ducibilities for 20 different metabolites that were contained in the standard mix and detected in the *E. coli* samples, it was clearly seen that they were about 2-fold smaller for the metabolites that were normalized by their corresponding stable isotopes as compared to the other analytes (Figure 24).

For the analysis of large sample batches, several quality control (QC) samples were interspersed throughout the measurement to ensure analytical stability.[83] If these QC samples failed to meet the established acceptance criteria [83, 126], measurement of the associated samples was repeated. These steps ensure that reliable quantitative data is generated and that the method can be applied over an extended period of time to large sample batches.
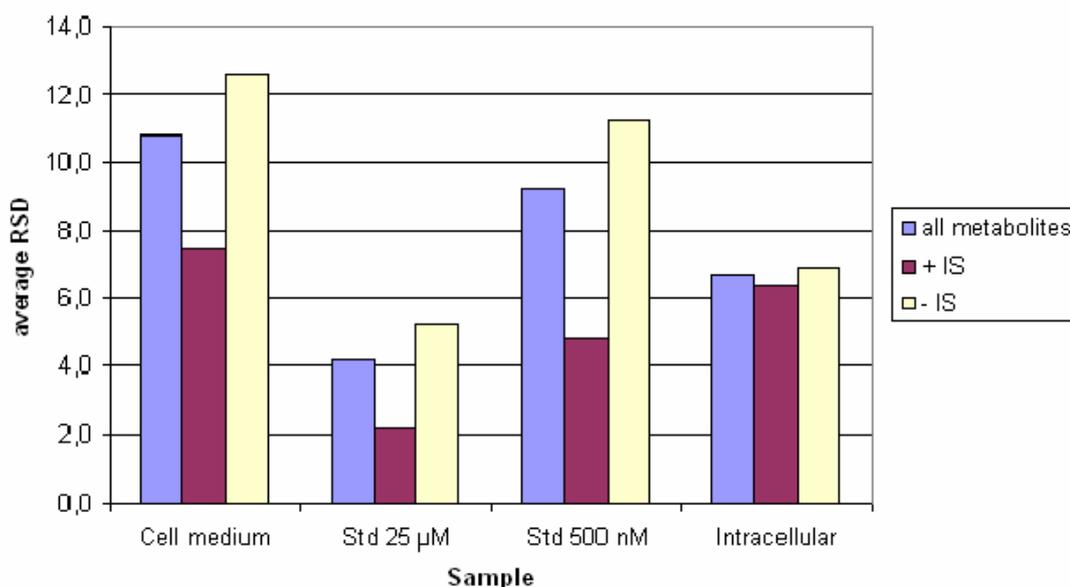
Figure 24: Intra-day reproducibility. Different samples were analyzed multiple times (n=5) over the course of 3 days. The average relative standard deviations (RSDs) for selected metabolites are shown. + IS: average RSD for metabolites with a matching stable isotope-labeled standard included. –IS: average RSD for metabolites that are normalized by a different standard.

Compared to methods based on AEC-MS/MS, the IP-LC-MS/MS approach has the advantage of detection limits that are 3-4 orders of magnitude lower.[13, 70] Therefore, more metabolites, including those that do not accumulate in the

cell, can be detected. While HILIC is well suited for the separation of polar analytes, average RSDs of 13% have been reported for the repeated analysis of cell extracts [5] and also in a direct comparison to RP-HPLC larger variabilities for peak areas have been observed.[131]

11 Analysis of glycolytic flux distribution in *E.coli*

11.1 Introduction

Flux analysis is a potent tool for the quantitative analysis of biochemical reaction networks, thus providing information on cellular regulation at the metabolic level.[13, 84, 132] Different strategies based on hyphenated mass spectrometry (MS) or nuclear magnetic resonance (NMR) have been used to analyze and quantify intracellular fluxes and metabolic networks.[85]

The most common approach consists of feeding the cells a labeled substrate, usually ^{13}C labeled glucose or glutamine, and then analyzing the label distribution, either in proteinogenic amino acids or in metabolic intermediates. The analysis of labeled amino acids is commonly done by GC-MS [86, 87, 133] after complete hydrolyzation of the intracellular proteins.

The second approach to metabolic flux analysis is the direct measurement of labeling distributions in intermediates of the central carbon metabolism. The analysis of flux distribution between glycolysis and PPP was performed by GC-MS measurement of lactate and ribose labeling, derived from $[1,2-^{13}C_2]$ glucose.[88, 89] Isotopomers of intracellular metabolites in *E. coli* were also measured by CE-TOF-MS [90] and compared with data from GC-MS analysis of proteinogenic amino acids. The correlation between the two methods was good, remaining differences were contributed to alterations in labeling distribution between precursors and amino acids as well as influences from flux modeling.[90] Another common method to analyze label distribution in selected intracellular metabolites is NMR.[134-136] The advantage of NMR in flux analysis is its ability to provide positional information on label distribution. However, high amounts of sample are required for ^{13}C-NMR approaches, which is the reason for increasing efforts being placed on the development of MS-based methods for flux analysis.

GC-MS based flux analysis raises several problems due to the need for derivatization, *i.e.* the addition of silica moieties, which contain a substantial fraction of natural isotopes [12, 86] and the high degree of fragmentation caused by electron impact ionization. Hence, a method that does not require derivatization but still yields information on positional distribution of isotopes would be of high value for the reliable acquisition of flux data. There are several reports on the use of LC-MS based methods in the context of flux analysis. An IP-LC-MS/MS based method was used to study the effect of a short glucose pulse in *E.coli* applying ultra-fast sampling.[92] Anion-exchange chromatography tandem mass spectrometry was used to evaluate mass isotopomer distributions of primary metabolites, which was subsequently fitted with a model to yield metabolic flux data.[93] A method for flux analysis in fed batch cultures by measuring intracellular free amino acids was described.[94] However, amino acids had to be derivatized prior to LC-MS/MS in order to increase their ionization efficiency.

Comprehensive flux analysis requires the modeling of all possible reactions encountered in the system of interest. However, it is often sufficient to roughly estimate changes in the relative flux of glucose into the direction of glycolysis and PPP. The second aim of this study was, thus, to develop and validate a method that allows a fast and simple evaluation of shifts in the upper part of the central carbon metabolism. We decided to use excreted lactate as a read-out, since it is produced by numerous organisms under many different conditions in sufficient amounts. To test our method we used *E. coli* strains with different rates of glycolysis vs. PPP [12] that have already been described in chapter 8. With this system we evaluated the ability of our approach to estimate the flux distribution between glycolysis and PPP from the analysis of extracellular lactate. In addition, we evaluated labeling distribution in other intermediates of the central carbon metabolism, in order to detect alterations in the TCA and related pathways.

11.2 Material and Methods

11.2.1 LC-MS/MS

LC-MS/MS was performed as described in the previous chapter. Additional SRM transitions were analyzed for all possible isotope distributions of selected compounds (Table 13). Care was taken to keep the scan time for each compound at or above 30 ms. In order to still record enough data points for each peak (cycle time < 1.5 s), the run was subdivided into 4 periods, which were selected on basis of the retention times of compounds of interest.

11.2.2 Bacterial strains and growth conditions

The *E. coli* wild type strain MG1655 and its mutant *PntAB-UdhA* were cultured at 37ºC in 250-mL shake flasks containing 30 mL minimal (M9) medium supplemented with 3 g/L glucose. An aliquot of an overnight culture corresponding to an OD of 0.1 in 30 mL (V1) was taken and centrifuged (5 min, 4500 rpm at 4°C) to separate the cells from the medium. The cell pellet was washed three times with 2 mL of glucose-free medium to remove all metabolites that had been secreted into the medium. The cells were then used to inoculate the shake flasks containing 30 mL M9 medium containing [1,2-^{13}C$_2$]glucose. Harvesting was performed in strictly exponential growth. For the control experiment, cells were also washed with M9 medium containing 3 g/L [1,2-^{13}C$_2$]glucose. The cells were then used to inoculate the 250 mL shake flasks containing 30 mL M9 medium with [1,2-^{13}C$_2$]glucose. Harvesting and extractions as well as quantification were also done as described above.

Table 13: SRM Transitions for selected metabolites used in flux analysis. Transitions indicated in bold were used as signature isotopomers

Analyte	Product ion	unlabeled		+1		+2		+3		+4		+5		+6	
		Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3	Q1	Q3
Glutamate	$-CO_2$	146	102	147	102/103	148	103/104	**149**	104/**105**	150	105/106	**151**	**106**		
Aspartate	$-CO_2$	132	88	133	88/89	134	89/90	**135**	**90**/91	**136**	**91**				
Lactate	$-CO_2$	**89**	**43**	90	43/44	**91**	**44/45**	92	45						
G6P	$[H_2PO_4]^-$	259	97	260	97	261	97	262	97	263	97	264	97	265	97
F6P	$[H_2PO_4]^-$	259	97	260	97	261	97	262	97	263	97	264	97	265	97
Glycerol-1P	$[PO_3]^-$	171	79	172	79	173	79	174	79						
Glycerol-2P	$[PO_3]^-$	171	79	172	79	173	79	174	79						
Pyruvate	$-CO_2$	**87**	**43**	88	43/44	**89**	**44/45**	89	45						
Succinate	$-CO_2$	117	73	118	73/74	119	74/75	**120**	**75/76**	**121**	**76**				
Malate	$-H_2O$	133	115	134	116	135	117	**136**	**118**	**137**	**119**				
2-Ketoglutarate	$-CO_2$	145	101	146	101/102	147	102/103	**148**	**103/104**	149	104/105	**150**	**105**		
Fumarate	$-CO_2$	115	71	116	71/72	117	72/73	**118**	**73/74**	119	74				
P-Glycerate	$[H_2PO_4]^-$	**185**	**97**	186	97	**187**	**97**	188	97						
FBP	$[H_2PO_4]^-$	338	97	339	97	340	97	341	97	342	97	343	97	344	97
cis-Aconitate	$-2\ CO_2$	173	85	174	85/86	175	85/86/87	**176**	**86**/87/88	177	87/88/89	178	89	**179**	**89**
DHAP	$[H_2PO_4]^-$	**169**	**97**	170	97	**171**	**97**	172	97						
Lactoylglutathione	Glutathione	**378**	**306**	379	306	**380**	**306**								
Acetyl-CoA	$[PO_3]^-$	808	79	809	79	810	79	811	79	812	79				

11.2.3 Experimental details for cell harvesting and metabolite extraction

Cell harvesting and extraction was done as described in Chapter 7, with the following modifications. After fast filtration, an aliquot of the filtered medium was taken for the analysis of extracellular metabolites. The extraction solution consisted of 80% (v/v) methanol in water at -20ºC, containing the extraction standards [2,3,3,3-^{2}H$_4$]lactate, [U-^{2}H]succinate, [2,3,3-^{2}H$_3$]malate, [2,3,3,4,4-^{2}H$_5$]glutamate, [U-^{13}C]fumarate, [U-^{13}C]pyruvate, [U-^{13}C]glucose-1-phosphate and [U-^{13}C, U-^{15}N]AMP. Sample preparation for media samples was very simple, a 100 µL aliquot was mixed with internal standard and directly measured by LC-MS/MS.

11.3 Results and Discussion

11.3.1 ^{13}C-Label distribution in extracellular lactate

For the measurement of lactate in cell medium, several milliliters of medium were collected after fast filtration of cells. An aliquot of 100 µL medium was mixed with internal standards and directly measured by LC-MS/MS. Lactate fragments under the chosen SRM parameters by loss of C1 as H_2CO_2 (Figure 25).
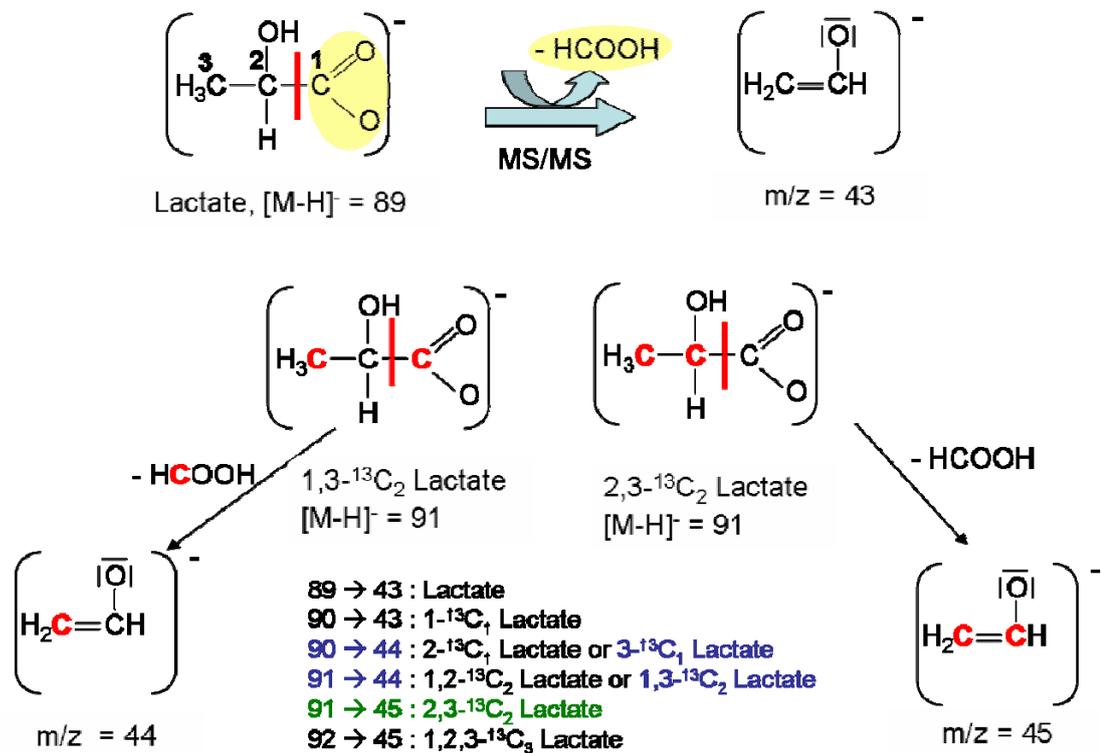
Figure 25: Fragmentation and labelling pattern of lactic acid in MS/MS experiments. Bold red fonts indicate ^{13}C labeled carbon atoms.

This fact can be used to distinguish different labeling patterns derived from metabolization of labeled glucose via different pathways. If $[1,2-^{13}C_2]$glucose is metabolized by glycolysis, the label is redistributed to C2 and C3 of lactate. On the other hand, in the PPP, $1,3-^{13}C_2$-lactate as well as well as $3-^{13}C_1$-lactate are produced (Figure 26). By comparing the excretion of the different isotopes, one can estimate the relative flux distributions of glucose between glycolysis and the PPP.
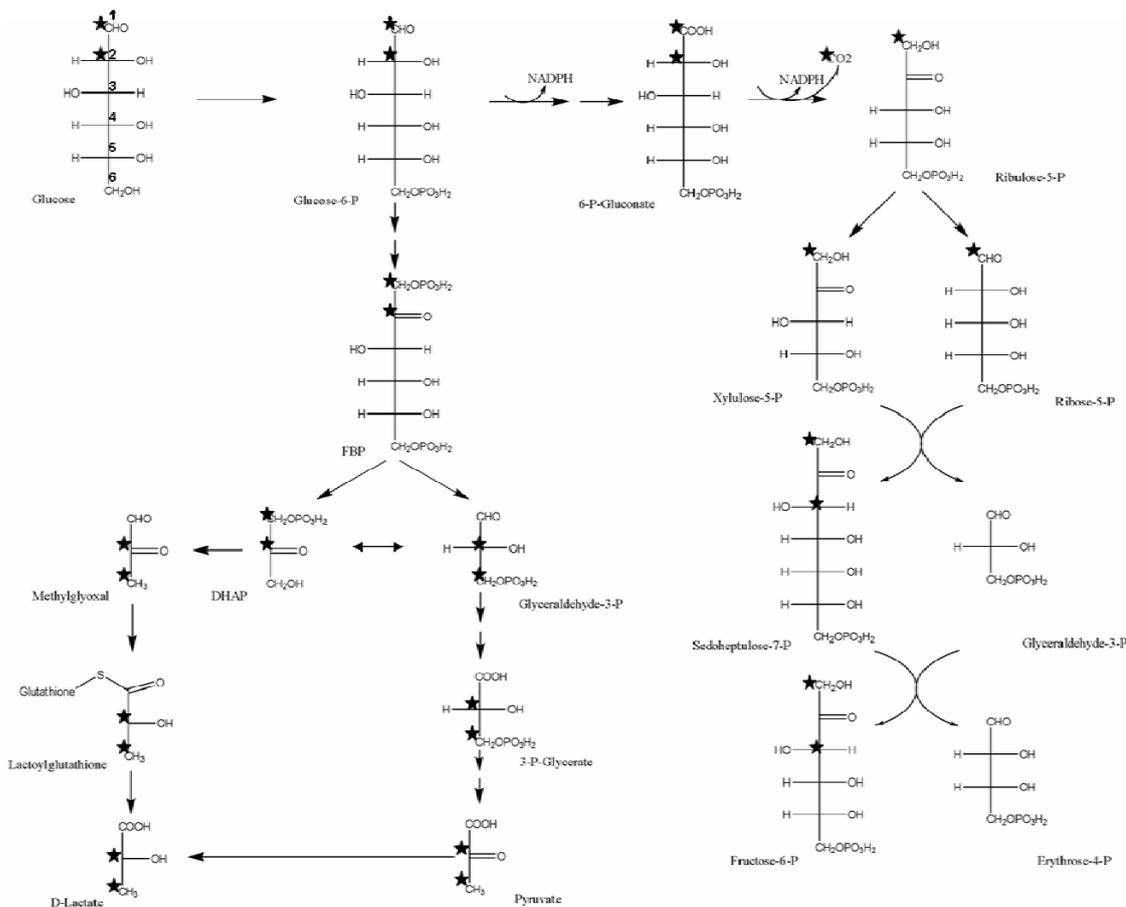
Figure 26: Metabolism of [1,2-^{13}C$_2$]glucose. Label distribution through important intermediates of glycolysis, PPP and MG-pathway is shown. ^{13}C labelled atoms are indicated by a star ($\star$).

In a proof-of-principle study, we compared the glucose metabolism of a double-deletion mutant of *E. coli,* which lacks the transhydrogenases UdhA and PntAB, with that of the corresponding wild type strain. The membrane-bound proton-translocating transhydrogenase PntAB is a major source of NADPH in *E. coli*, that produces during aerobic batch growth 35–45% of the NADPH required for biosynthesis, whereas pentose phosphate pathway and isocitrate dehydrogenase contribute 35–45% and 20–25%, respectively.[12] The soluble energy-independent transhydrogenase UdhA, in contrast, catalyzes the reoxidation of excess NADPH. Consequently, deletion of these two transhydrogenases will lead to a compensatory increase in the generation of anabolic reductant NADPH in the first reaction of the oxidative branch of the PPP, the dehydrogenation of G6P to 6-phosphoglucono-δ-lactone, and by

isocitrate dehydrogenase. The increased flux through the PPP will in turn result in a decreased glucose flux through glycolysis as already demonstrated by Sauer et al. (2004) employing metabolic flux ratio analysis by GC-MS. It was also demonstrated that metabolite levels for many compounds of the central carbon metabolism are significantly different in the mutant as compared to the wild type strain.[137] In the present study, our aim was to evaluate the ability of the RR-IP-LC-MS/MS method to detect these changes in glycolytic fluxes based on the isotope distribution in extracellular lactic acid.

To this end the two strains were grown on minimal medium containing 100% [1,2-^{13}C$_2$]glucose. In initial experiments it became obvious that the aliquot of the over-night culture used to inoculate the shake flasks contained a high amount of unlabeled lactate (data not shown). Since we were only interested in lactate excreted by the cells under our experimental conditions, we decided to wash the inoculum with fresh, glucose-free medium before use. We measured the supernatant after each washing step and found that all the extracellular lactate was removed after the third iteration.

As described above, we used the ratio of 2,3-^{13}C$_2$-lactate in the wild type *vs.* the mutant strain to estimate the difference in flux through glycolysis. Published data show that the flux through glycolysis was 68% of total glucose flux in the wild type strain,[138] while it was estimated at 52% in the *PntAB-UdhA* mutant.[12] This amounts to a ratio of 0.76 (PU vs. wt). Our measurement yielded a ratio of 0.71. For the flux through the PPP, values of 32% (wt) and 47% (PU) had been reported.[12, 138] The theoretical ratio of 1.49 (PU vs. wt) again corresponds well with our measured value of 1.32 (Figure 27). Thus, our method is clearly able to give a first estimate on alterations in flux distributions, without the need for elaborate modeling and computation. However, it is not feasible to calculate absolute flux values, at least not without prior knowledge of absolute flux values in a reference strain, in part because lactate from glycolysis is found in much higher amounts than lactate from PPP.
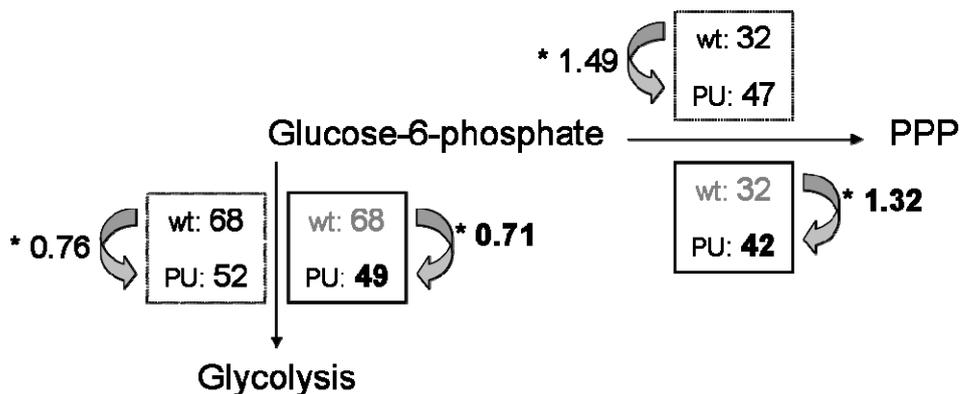
Figure 27: Flux distribution in the wild type strain as compared to the *PntAB-UdhA* mutant. Bold entries in the solid boxes denote our estimates in comparison to the estimates reported by [12, 138], respectively, which are listed in the dashed boxes.

11.3.2   The methylglyoxal pathway

As shown in Figure 26, in the course of glycolysis FBP is split in a reversible enzymatic reaction catalyzed by FBP-aldolase into DHAP and glyceraldehyde-3-P. Therefore, one would expect a ratio of unlabeled vs. +2-labelled lactate of 1 (i.e. 50% unlabelled vs. 50% labeled). After we had started washing the inoculum to remove any extracellular metabolites, especially unlabeled lactate, we consistently found an excess of 2,3-^{13}C$_2$-lactate as compared to unlabeled lactate and this difference was always higher for the wild type (~0.8) as for the mutant (~0.9). The same observation was made for several intracellular metabolites such as pyruvate or DHAP, yet not for 3-phosphoglycerate (Figure 28).

G6P

↓
↓

FBP

wt: 0.83
PU: 0.95

DHAP ⟷ G3P

Methylglyoxal

P-Glycerate

wt: 0.98
PU: 1.13

Lactoylglutathione

wt: 0.76
PU: 0.92

Lactate ⟷ Pyruvate

wt: 0.87
PU: 0.97
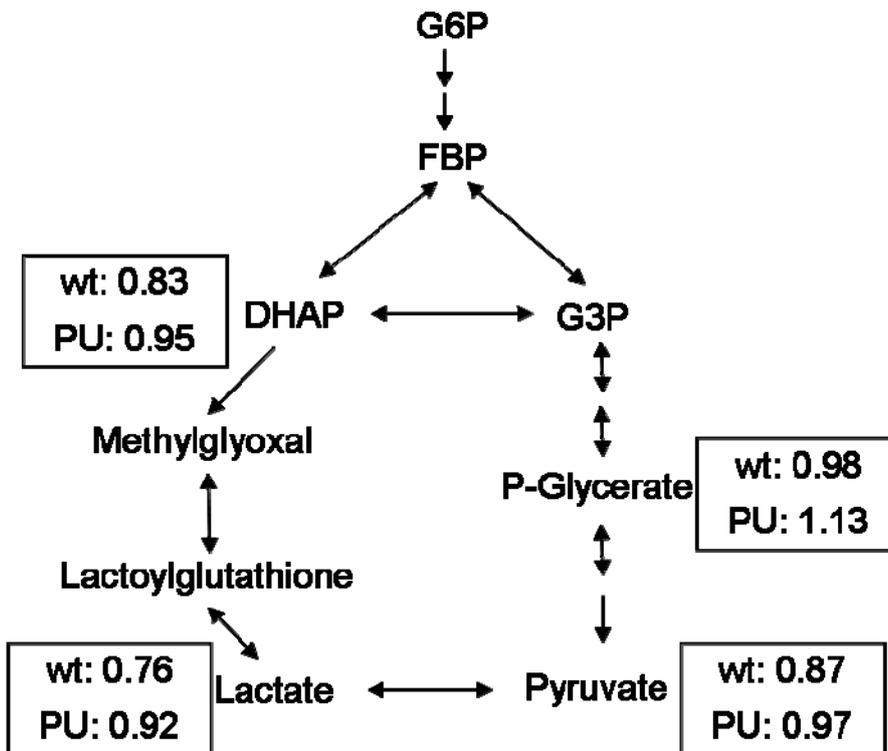
Figure 28: Ratios of the unlabelled vs. the +2-labelled isotopomer for different compounds. The upper entry denotes the ratio for the wild type strain, while the lower entry indicates the value for the *PntAB-UdhA* mutant.

After excluding measurement errors and ensuring the purity of the labeled glucose, we investigated the only step during glycolysis, where the labeled half of the $[1,2-^{13}C_2]$glucose molecule reacts differently than the lower, unlabeled half, i.e. the conversion of FBP to DHAP and GAP. The upper three carbon atoms, C1 – C3, of FBP initially end up in DHAP, while C4 – C6 end up in GAP. The enzyme triosephosphate isomerase, Tpi, catalyzes the isomerization of these two compounds. In Δtpi strains, the accumulating DHAP is processed by a different set of reactions, the methylglyoxal (MG) pathway (Figure 26).[138] DHAP is converted to MG by methylglyoxal synthase. MG is toxic for *E. coli* cells and, therefore, is removed by conversion to D-lactate via lactoylglutathione or L-lactate, which requires NADPH.[139]

However, MG pathway activity has been described in other *E. coli* strains; it was identified as a source of chiral contamination in lactate and alanine fermentation.[139, 140] The pathway was also active after a glucose pulse,

presumably to replenish the intracellular inorganic phosphate pool and to prevent the accumulation of phosphorylated intermediates of glycolysis.[141] We propose that MG pathway activity has been responsible for the imbalance in the ratio of unlabeled to labeled lactate, by directly shuttling DHAP towards lactate and pyruvate, respectively. This would explain the observed results for lactate, pyruvate and DHAP and the different result for phosphoglycerate. In order to test this hypothesis we measured a specific intermediate of the MG pathway, lactoylglutathione (LG). The specific SRM transition 378 → 306 (loss of the lactate moiety) was chosen for the measurement of intracellular LG. A clear LG peak was seen in all the samples derived from the washed inoculum, while no LG was seen in control samples grown on LB medium (Figure 29). Also, about 2.5 times more LG was detected in the wild type samples as compared to the mutant and the ratio of LG vs. LG+2 was 95% for the mutant and 89% for the wild type strain. These findings strongly support the hypothesis that the MG pathway was upregulated under the applied conditions.
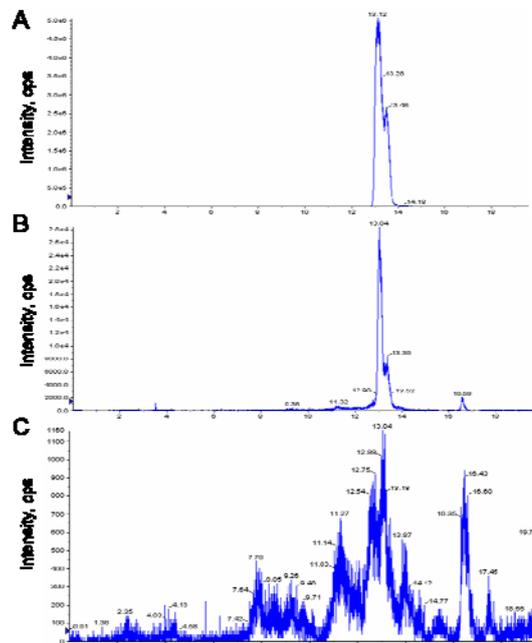
Figure 29: Measurement of LG in different samples. A) LG standard in water, 10 µM. B) *E. coli* sample grown on M9 medium, washed inoculum. C) *E. coli* sample grown on LB medium.

As mentioned above, the activation of the MG pathway is described under conditions of glucose limitation and subsequent glucose pulse.[141] We therefore wanted to evaluate the influence of washing the inoculum on the observed flux distribution. To this end we repeated the experiment as described above, with the exception of washing the inoculum with M9 medium supplemented with 3 g/L [1,2-$^{13}C_2$]glucose. This was done in order to maintain the extracellular glucose level, while at the same time preventing the formation of unlabelled glucose metabolites. Under these conditions, we found ratios close to 1 for the unlabelled vs. the +2-labelled isotopomers of the metabolites involved in the lower part of glycolysis (Figure 30). Thus, the initial washing procedure is substantially responsible for the observed flux through the MG pathway. In order to definitely exclude any effect of the MG pathway on the label distribution towards lactate, an MG pathway deletion mutant would have to be analyzed.
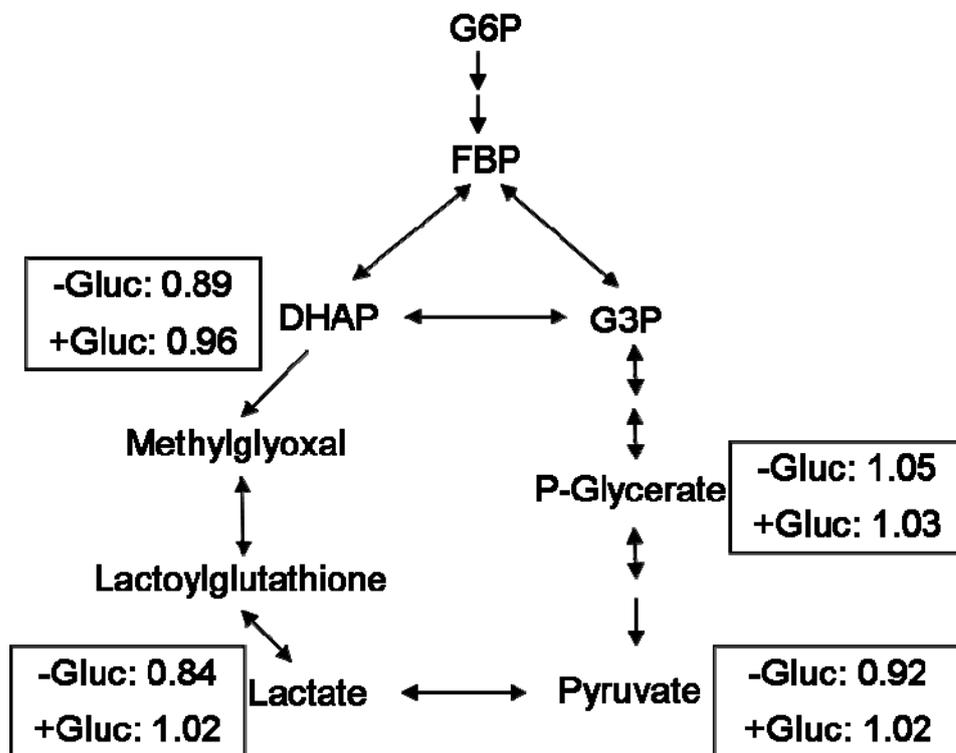

Figure 30: Ratios of the unlabelled vs. the +2-labelled isotopomer for different compounds. The upper entry denotes the ratio (average of the two strains) upon glucose deprivation during sample preparation while the lower entry denotes the ratio upon addition of labelled glucose during the washing step.

99

11.3.3 Information derived from other metabolites – the tricarboxylic acid cycle (TCA)

The findings for the different isotopomers of lactate described in the previous sections are very well suited for the rapid assessment of perturbations in the upper part of the central carbon metabolism. However, additional information on cellular processes can be gained from the analysis of other metabolites and their corresponding isotopes. Nevertheless, since most metabolites are not excreted into the cell medium in amounts similar to lactic acid, they are preferably analyzed in cell extracts. We established SRM transitions for all possible isotopomers of the most informative intermediates of the central carbon metabolism that were amenable to IP-LC-MS/MS analysis (Table 13). These measurements yielded a substantial amount of data, in part due to the fact that different possibilities exist for loss or retention of the label during fragmentation in MS/MS mode. Also, the multitude of analyzed SRMs necessitated the subdivision of the MS run into four periods. However, this fact could also be used to gain additional information on the localization of labeled ^{13}C atoms in the molecule, which was useful in the correlation of isotope labeling patterns with their corresponding metabolic pathways.

During the analysis of isotopomer distributions, remarkable differences for different labeling states in the wild type vs. the *PntAB-UdhA* mutant strain became apparent. As depicted in Figure 31, significant differences, as determined by a Student's t-test with P<0.05, were for example seen for the +3 and the fully labeled states in the two strains. While the +3-labeling state was always higher in the mutant strain, the uniformly labeled state was found to a higher degree in the wild type strain.
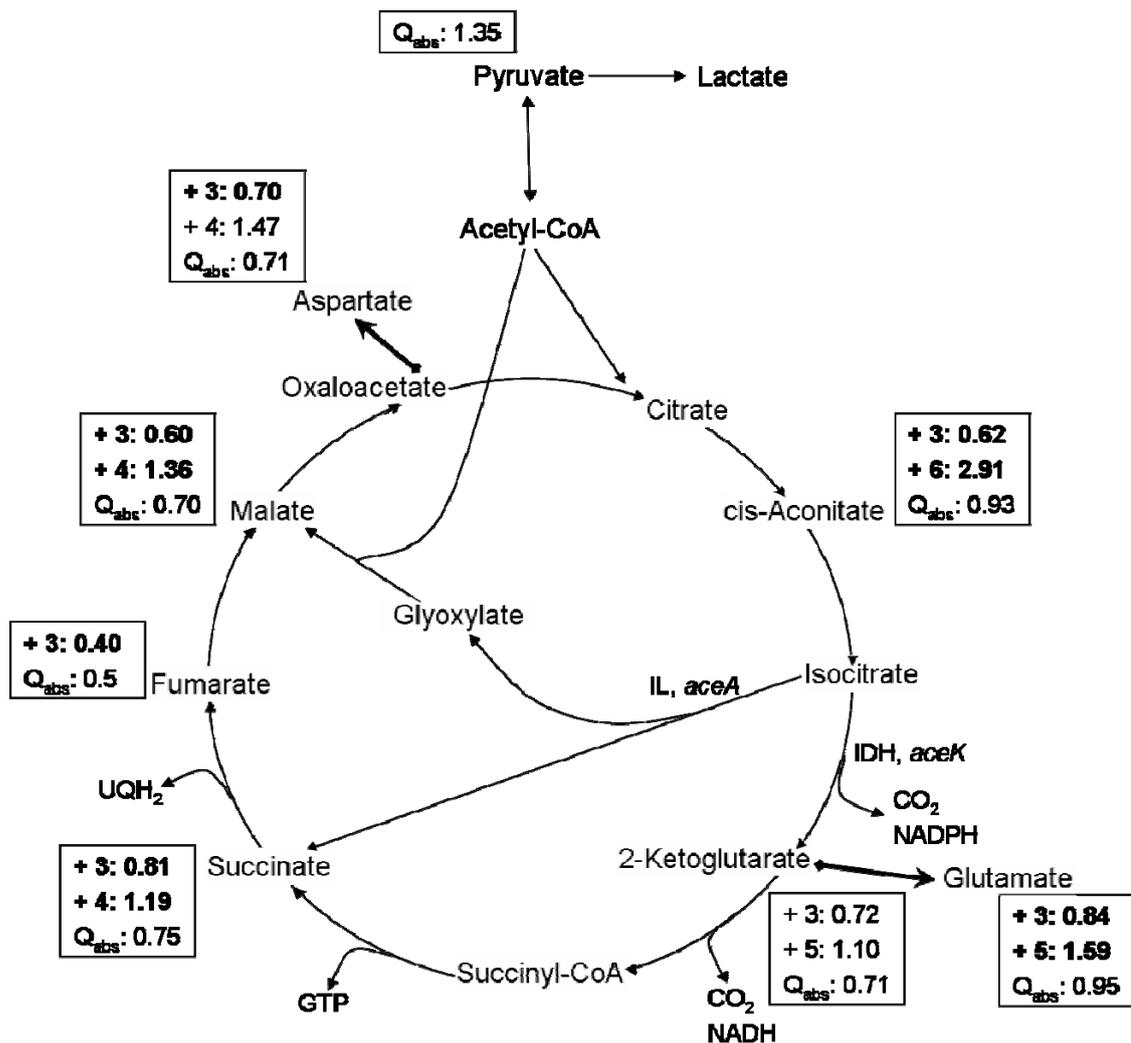
Figure 31: Ratio of two different labelling states (+3 vs. fully labeled) for different TCA intermediates in the wild type vs. the *PntAB-UdhA* mutant strain and ratios of absolute metabolites levels (Q_{abs}) in the wild type vs. the mutant strain. Values marked in bold indicate significant differences between the different labelling states of the two strains (Student's t-test, $p<0.05$).

We then investigated in more detail the ^{13}C-label distribution in the TCA cycle. In the first pass through the TCA cycle, both ^{13}C-labeled atoms from acetyl-CoA are retained in the metabolic intermediates (Figure 32 A/I). Succinate is a symmetric molecule; therefore, the labeling pattern of the subsequent metabolites is not fixed. However, for the conclusions drawn here, it does not make any difference, if these metabolites are labeled at positions 1 and 2 or at positions 3 and 4.
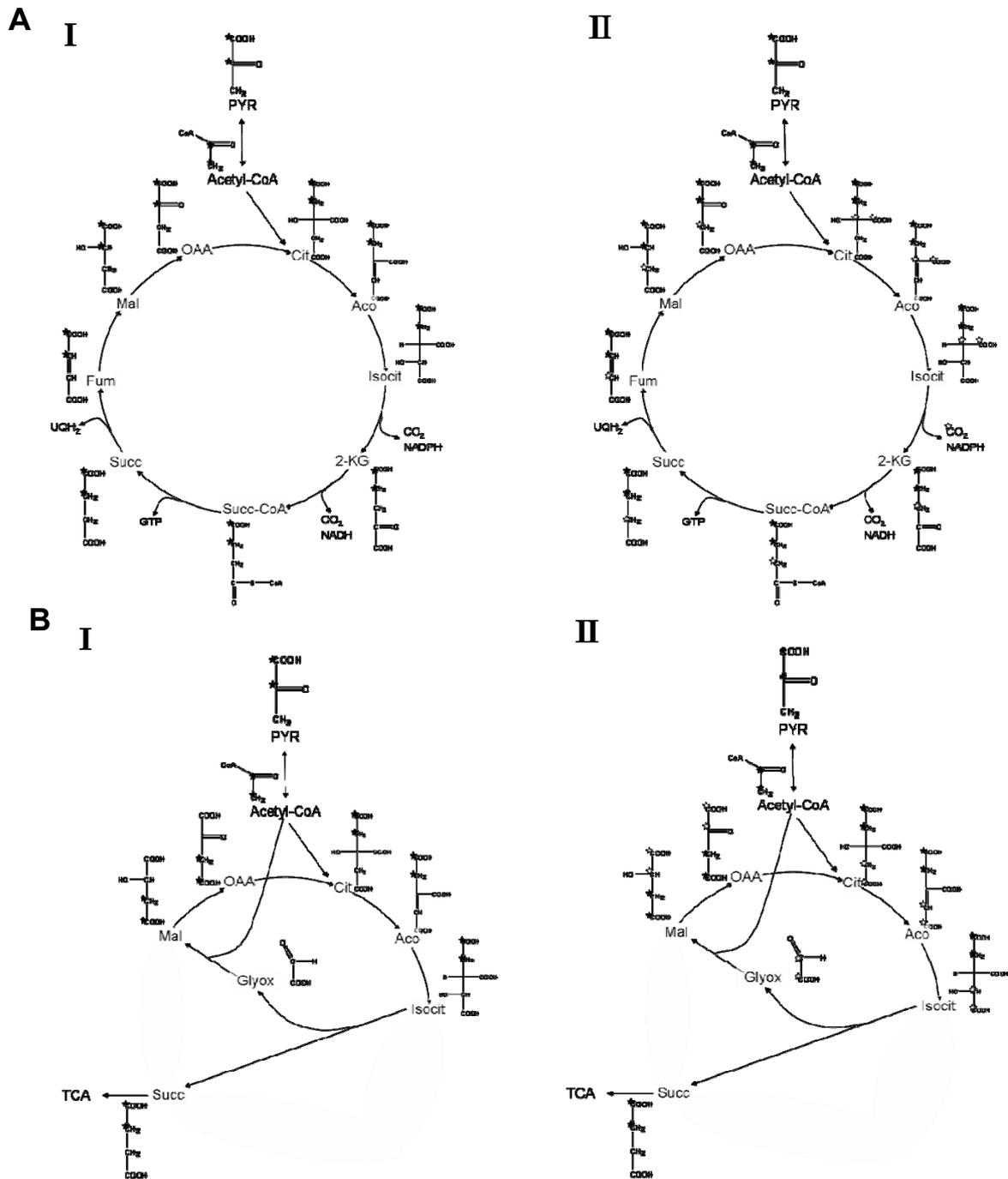
Figure 32: Distribution of ^{13}C-labeling in the first (I) and the second (II) pass through the tricarboxylic acid cycle (A) and the glyoxylate pathway (B). Solid black stars (★) denote labelled ^{13}C atoms entering the cycle. Empty stars (☆) denote labelled atoms in that had entered from the previous pass. Succinate is a symmetric molecule and, therefore, it is impossible to discern whether positions 1 and 2 or 3 and 4 are isotope labeled. For clarity's sake, only one of the two possibilities is denoted. In both cases one labelled $^{13}CO_2$ molecule would leave in the next turn of the TCA.

This is due to the fact that in the second pass through the TCA (Figure 32 A/II), one of the two labeled ^{13}C atoms is lost as $^{13}CO_2$, leaving a molecule with a uneven number of labeled carbons. The higher number of +3 labeled intermediates in the mutant strain might thus be explained by the higher flux through the TCA, as was already hypothesized previously.[83]

However, this did not explain the opposite trend for fully labeled intermediates. Therefore, we investigated alterations in the TCA cycle and found several descriptions on the activity of the glyoxylate shunt as a bypass flux of the TCA.[138] While this pathway is mainly activated upon growth on acetate [142], it was also described under conditions of glucose limitation [143, 144] and as modulated by the NADPH/NADP ratio.[145] Therefore, we hypothesized that the glyoxylate shunt is activated upon glucose limitation during the initial washing step, and that this activation is substantially larger in the wild type strain than in the mutant. This could be due to the fact that the *PntAB-UdhA* mutant has to rely on the activity of the enzyme isocitrate dehydrogenase in order to produce NADPH, while the wild type strain uses the bypass to keep up the production of amino acid precursors while at the same time balancing its NADPH/NADP ratio. The distribution of ^{13}C-labeled atoms through the glyoxylate bypass is depicted in Figure 32/B. It is evident, that the activity of this metabolic pathway would lead to an increased production of fully labeled intermediates, as observed predominantly in the wild type strain (Figure 31). We tried to confirm this hypothesis by measuring intracellular levels of glyoxylate. However, ionization efficiency of the main product was insufficient for detection of glyoxylate in our *E. coli* samples.

As noted above, the measurement of all possible isotopomers of even a selection of analytes leads to a high number of SRM transitions. Since dwell times have to be set to at least 30 ms and a sufficient number of data points have to be recorded for each peak for reliable quantification, this would necessitate multiple analyses for each sample. To circumvent this severe drawback towards high-throughput analysis, we established specific SRMs from the observations described above without limiting the drawing of

meaningful conclusions on carbon fluxes from the data. This signature SRMs, as depicted in Table 13 in bold, can be easily measured in a single run, which is subdivided into four periods, with a maximum of 17 SRMs per period, corresponding to a cycle time of ~0.5 s.

We also measured the absolute intracellular levels of several TCA cycle inter-mediates and the amino acids aspartate and glutamate by summing all detected isotopomers for each compound. The results obtained correlated well with our previously published CE-TOF-MS data; for instance, the levels of fumarate measured by LC-MS/MS were 0.27 and 0.54 nmol per mg cell dry weight (cdw) for wild type and mutant, respectively, while the corresponding CE-TOF-MS values were 0.35 and 0.54.[18] Generally, higher amounts of TCA cycle intermediates were found for the mutant strain as compared to the wild type (Figure 31).

12 Conclusion and Outlook

To date the most promising approach to tackle the enormous complexity of the metabolome is the combination of several different methods for metabolome analysis. In this thesis I demonstrated the development, validation and application of two different methodologies. Each of them is highly suitable to certain biological problems but can also be used as more universal analytical tools. The strength of the CE-TOF-MS based method is its ability for high-confidence identification of additional metabolites of interest, while still providing reliable quantitative data. The RR-IP-LC-MS/MS method yields accurate quantitative information on selected metabolites over a wide dynamic range and is very well suited for the highly sensitive detection of minor components of a complex mixture. Nevertheless, it can also be used for initial screening approaches that can later be transferred to a high-resolution MS system for identification.

The methods were applied to various biological problems, with a high emphasis on thorough validation of the complete sample preparation and measurement process. This was achieved by optimizing protocols and by the introduction of several internal standards as well as different quality control measures. We could demonstrate that our approaches were able to reliably answer a variety of biological questions in different cellular systems.

However, several points remain to be addressed:

- The panel of metabolites that are absolutely quantified by both methods has to be extended, as well as the number of confirmed metabolite hits derived from TOF-MS based analysis

- Along the same line the panel of internal standard compounds should be extended. A promising approach in this regard is the use of uniformly ^{13}C, ^{15}N labeled biomass, which theoretically would provide standards for all the metabolites contained in the organism at much lower overall costs. The suitability of this procedure is now under investigation.

- Regarding the data on flux analysis in *E. coli*, additional strains with disruptions in the central carbon metabolism have to be analyzed in order to evaluate the ability of the described method to detect these alterations in a reliable fashion. Also, the kinetic of label distribution throughout the cellular network is of high interest, as well as its time-resolved behavior to perturbations. In order to cope with the large number of samples, analysis time will have to be reduced further by using even smaller stationary phase particles and higher column pressures.

13 References

(1)   Oliver, S. G.; Winson, M. K.; Kell, D. B.; Baganz, F. *Trends Biotechnol* **1998**, *16*, 373-378.

(2)   Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrom Rev* **2006**.

(3)   Tolstikov, V. V.; Fiehn, O. *Anal Biochem* **2002**, *301*, 298-307.

(4)   Lafaye, A.; Labarre, J.; Tabet, J. C.; Ezan, E.; Junot, C. *Anal Chem* **2005**, *77*, 2026-2033.

(5)   Bajad, S. U.; Lu, W.; Kimball, E. H.; Yuan, J.; Peterson, C.; Rabinowitz, J. D. *J Chromatogr A* **2006**, *1125*, 76-88.

(6)   Fiehn, O.; Kopka, J.; Dormann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L. *Nat Biotechnol* **2000**, *18*, 1157-1161.

(7)   Bolling, C.; Fiehn, O. *Plant Physiol* **2005**, *139*, 1995-2005.

(8)   Nikiforova, V. J.; Kopka, J.; Tolstikov, V.; Fiehn, O.; Hopkins, L.; Hawkesford, M. J.; Hesse, H.; Hoefgen, R. *Plant Physiol* **2005**, *138*, 304-318.

(9)   Roessner, U.; Wagner, C.; Kopka, J.; Trethewey, R. N.; Willmitzer, L. *Plant J* **2000**, *23*, 131-142.

(10)   Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T. *J Proteome Res* **2003**, *2*, 488-494.

(11)   Luo, B.; Groenke, K.; Takors, R.; Wandrey, C.; Oldiges, M. *J Chromatogr A* **2007**, *1147*, 153-164.

(12)   Sauer, U.; Canonaco, F.; Heri, S.; Perrenoud, A.; Fischer, E. *J Biol Chem* **2004**, *279*, 6613-6619.

(13)   Timischl, B.; Dettmer, K.; Oefner, P. J. *Anal Bioanal Chem* **2008**, *391*, 895-898.

(14)   Weckwerth, W. *Annu Rev Plant Biol* **2003**, *54*, 669-689.

(15)   Villas-Boas, S. G.; Mas, S.; Akesson, M.; Smedsgaard, J.; Nielsen, J. *Mass Spectrom Rev* **2005**, *24*, 613-646.

(16)   Dettmer, K.; Hammock, B. D. *Environ Health Perspect* **2004**, *112*, A396-397.

(17)   Fiehn, O. *Plant Mol Biol* **2002**, *48*, 155-171.

(18)   Allen, J.; Davey, H. M.; Broadhurst, D.; Heald, J. K.; Rowland, J. J.; Oliver, S. G.; Kell, D. B. *Nat Biotechnol* **2003**, *21*, 692-696.

(19)   Heiger, D. *High performance capillary electrophoresis - An introduction*: Germany, 2000.

(20) Cornish, T. J.; Cotter, R. J. *Rapid Commun Mass Spectrom* **1993**, *7*, 1037-1040.

(21) McIntyre, D. In *Technical Overview*; Technologies, A., Ed., 2004.

(22) Want, E. J.; Nordstrom, A.; Morita, H.; Siuzdak, G. *J Proteome Res* **2007**, *6*, 459-468.

(23) Kind, T.; Fiehn, O. *BMC Bioinformatics* **2006**, *7*, 234.

(24) Ojanpera, S.; Pelander, A.; Pelzing, M.; Krebs, I.; Vuori, E.; Ojanpera, I. *Rapid Commun Mass Spectrom* **2006**, *20*, 1161-1167.

(25) Soga, T.; Ueno, Y.; Naraoka, H.; Ohashi, Y.; Tomita, M.; Nishioka, T. *Anal Chem* **2002**, *74*, 2233-2239.

(26) Stolker, A. L.; Niesing, W.; Fuchs, R.; Vreeken, R. J.; Niessen, W. M.; Brinkman, U. A. *Anal Bioanal Chem* **2004**, *378*, 1754-1761.

(27) Smith, B. *LC/MS Applications Booklet* St-Quentin-en-Yvelines Cedex, 2001.

(28) Klapa, M. I.; Aon, J. C.; Stephanopoulos, G. *Biotechniques* **2003**, *34*, 832-836, 838, 840 passim.

(29) Douglas, D. J.; Frank, A. J.; Mao, D. *Mass Spectrom Rev* **2005**, *24*, 1-29.

(30) Xia, Y. Q.; Miller, J. D.; Bakhtiar, R.; Franklin, R. B.; Liu, D. Q. *Rapid Commun Mass Spectrom* **2003**, *17*, 1137-1145.

(31) Le Blanc, J. C.; Hager, J. W.; Ilisiu, A. M.; Hunter, C.; Zhong, F.; Chu, I. *Proteomics* **2003**, *3*, 859-869.

(32) Belov, M. E.; Anderson, G. A.; Angell, N. H.; Shen, Y.; Tolic, N.; Udseth, H. R.; Smith, R. D. *Anal Chem* **2001**, *73*, 5052-5060.

(33) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom Rev* **1998**, *17*, 1-35.

(34) Sancho, J. V.; Pozo, O. J.; Ibanez, M.; Hernandez, F. *Anal Bioanal Chem* **2006**, *386*, 987-997.

(35) Gentili, A.; Sergi, M.; Perret, D.; Marchese, S.; Curini, R.; Lisandrin, S. *Rapid Commun Mass Spectrom* **2006**, *20*, 1845-1854.

(36) Eisenreich, W.; Bacher, A. *Phytochemistry* **2007**, *68*, 2799-2815.

(37) Slupsky, C. M.; Rankin, K. N.; Wagner, J.; Fu, H.; Chang, D.; Weljie, A. M.; Saude, E. J.; Lix, B.; Adamko, D. J.; Shah, S.; Greiner, R.; Sykes, B. D.; Marrie, T. J. *Anal Chem* **2007**, *79*, 6995-7004.

(38) Lewis, I. A.; Schommer, S. C.; Hodis, B.; Robb, K. A.; Tonelli, M.; Westler, W. M.; Sussman, M. R.; Markley, J. L. *Anal Chem* **2007**, *79*, 9385-9390.

(39) Rezzi, S.; Ramadan, Z.; Fay, L. B.; Kochhar, S. *J Proteome Res* **2007**, *6*, 513-525.

(40) Weljie, A. M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C. M. *Anal Chem* **2006**, *78*, 4430-4442.

(41) Dunn, W. B.; Bailey, N. J.; Johnson, H. E. *Analyst* **2005**, *130*, 606-625.

(42) Mims, D.; Hercules, D. *Anal Bioanal Chem* **2003**, *375*, 609-616.

(43) Mims, D.; Hercules, D. *Anal Bioanal Chem* **2004**, *378*, 1322-1326.

(44) Edwards, J. L.; Kennedy, R. T. *Anal Chem* **2005**, *77*, 2201-2209.

(45) Vaidyanathan, S.; Jones, D.; Ellis, J.; Jenkins, T.; Chong, C.; Anderson, M.; Goodacre, R. *Rapid Commun Mass Spectrom* **2007**, *21*, 2157-2166.

(46) Vaidyanathan, S.; Goodacre, R. *Rapid Commun Mass Spectrom* **2007**, *21*, 2072-2078.

(47) Mala, Z.; Krivankova, L.; Gebauer, P.; Bocek, P. *Electrophoresis* **2007**, *28*, 243-253.

(48) Shihabi, Z. K. *Electrophoresis* **2000**, *21*, 2872-2878.

(49) Soga, T.; Baran, R.; Suematsu, M.; Ueno, Y.; Ikeda, S.; Sakurakawa, T.; Kakazu, Y.; Ishikawa, T.; Robert, M.; Nishioka, T.; Tomita, M. *J Biol Chem* **2006**, *281*, 16768-16776.

(50) Soga, T.; Kakazu, Y.; Robert, M.; Tomita, M.; Nishioka, T. *Electrophoresis* **2004**, *25*, 1964-1972.

(51) Majors, R. E., and Przybyciel, M. *LCGC* **2002**, *20*, 584-593.

(52) Ishizuka, N.; Kobayashi, H.; Minakuchi, H.; Nakanishi, K.; Hirao, K.; Hosoya, K.; Ikegami, T.; Tanaka, N. *J Chromatogr A* **2002**, *960*, 85-96.

(53) Cabrera, K. *J Sep Sci* **2004**, *27*, 843-852.

(54) Xu, R. N.; Fan, L.; Rieser, M. J.; El-Shourbagy, T. A. *J Pharm Biomed Anal* **2007**, *44*, 342-355.

(55) Tanaka, N.; Kimura, H.; Tokuda, D.; Hosoya, K.; Ikegami, T.; Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Shintani, Y.; Furuno, M.; Cabrera, K. *Anal Chem* **2004**, *76*, 1273-1281.

(56) Ross, P.; Knox, J. H. *Adv Chromatogr* **1997**, *37*, 121-162.

(57) Buchholz, A.; Takors, R.; Wandrey, C. *Anal Biochem* **2001**, *295*, 129-137.

(58) Xing, J.; Apedo, A.; Tymiak, A.; Zhao, N. *Rapid Commun Mass Spectrom* **2004**, *18*, 1599-1606.

(59) Antonio, C.; Larson, T.; Gilday, A.; Graham, I.; Bergstrom, E.; Thomas-Oates, J. *J Chromatogr A* **2007**, *1172*, 170-178.

(60) Halász, I.; Endele, R.; Assauer, J. *J Chromatogr* **1975**, *112*, 37-60.

(61) de Villiers, A.; Lestremau, F.; Szucs, R.; Gelebart, S.; David, F.; Sandra, P. *J Chromatogr A* **2006**, *1127*, 60-69.

(62) Wilson, I. D.; Nicholson, J. K.; Castro-Perez, J.; Granger, J. H.; Johnson, K. A.; Smith, B. W.; Plumb, R. S. *J Proteome Res* **2005**, *4*, 591-598.

(63) Nordstrom, A.; O'Maille, G.; Qin, C.; Siuzdak, G. *Anal Chem* **2006**, *78*, 3289-3295.

(64) Stahlberg, J. *J Chromatogr A* **1999**, *855*, 3-55.

(65) Dai, J.; Carr, P. W. *J Chromatogr A* **2005**, *1072*, 169-184.

(66) Coulier, L.; Bas, R.; Jespersen, S.; Verheij, E.; van der Werf, M. J.; Hankemeier, T. *Anal Chem* **2006**, *78*, 6573-6582.

(67) Wamelink, M. M.; Struys, E. A.; Huck, J. H.; Roos, B.; van der Knaap, M. S.; Jakobs, C.; Verhoeven, N. M. *J Chromatogr B Analyt Technol Biomed Life Sci* **2005**, *823*, 18-25.

(68) Alpert, A. J.; Shukla, M.; Shukla, A. K.; Zieske, L. R.; Yuen, S. W.; Ferguson, M. A.; Mehlert, A.; Pauly, M.; Orlando, R. *J Chromatogr A* **1994**, *676*, 191-122.

(69) Brauer, M. J.; Yuan, J.; Bennett, B. D.; Lu, W.; Kimball, E.; Botstein, D.; Rabinowitz, J. D. *Proc Natl Acad Sci U S A* **2006**, *103*, 19302-19307.

(70) van Dam, J. C., Eman, M. R., Frank, J., Lange, H. C., van Dedem, G. W. K., Heijnen, S. J. *Analytica Chimica Acta* **2002**, *460*, 209-218.

(71) Wittmann, C.; Hans, M.; van Winden, W. A.; Ras, C.; Heijnen, J. J. *Biotechnol Bioeng* **2005**, *89*, 839-847.

(72) Visser, D.; van Zuylen, G. A.; van Dam, J. C.; Oudshoorn, A.; Eman, M. R.; Ras, C.; van Gulik, W. M.; Frank, J.; van Dedem, G. W.; Heijnen, J. J. *Biotechnol Bioeng* **2002**, *79*, 674-681.

(73) Lange, H. C.; Eman, M.; van Zuijlen, G.; Visser, D.; van Dam, J. C.; Frank, J.; de Mattos, M. J.; Heijnen, J. J. *Biotechnol Bioeng* **2001**, *75*, 406-415.

(74) Koek, M. M.; Muilwijk, B.; van der Werf, M. J.; Hankemeier, T. *Anal Chem* **2006**, *78*, 1272-1281.

(75) Villas-Boas, S. G.; Hojer-Pedersen, J.; Akesson, M.; Smedsgaard, J.; Nielsen, J. *Yeast* **2005**, *22*, 1155-1169.

(76) Denkert, C.; Budczies, J.; Kind, T.; Weichert, W.; Tablack, P.; Sehouli, J.; Niesporek, S.; Konsgen, D.; Dietel, M.; Fiehn, O. *Cancer Res* **2006**, *66*, 10795-10804.

(77) Phillips, J. B.; Beens, J. *J Chromatogr A* **1999**, *856*, 331-347.

(78) Ong, R. C.; Marriott, P. J. *J Chromatogr Sci* **2002**, *40*, 276-291.

(79) Pierce, K. M.; Hoggard, J. C.; Mohler, R. E.; Synovec, R. E. *J Chromatogr A* **2007**, *1184*, 341-352.

(80) Mohler, R. E.; Dombek, K. M.; Hoggard, J. C.; Pierce, K. M.; Young, E. T.; Synovec, R. E. *Analyst* **2007**, *132*, 756-767.

(81)   Guo, X.; Lidstrom, M. E. *Biotechnol Bioeng* **2007**, *99*, 929-940.

(82)   Shellie, R. A.; Welthagen, W.; Zrostlikova, J.; Spranger, J.; Ristow, M.; Fiehn, O.; Zimmermann, R. *J Chromatogr A* **2005**, *1086*, 83-90.

(83)   Timischl, B.; Dettmer, K.; Kaspar, H.; Thieme, M.; Oefner, P. J. *Electrophoresis* **2008**, *29*, 2203-2214.

(84)   Sauer, U. *Curr Opin Biotechnol* **2004**, *15*, 58-63.

(85)   Zamboni, N. *Toward metabolome-based ^{13}C flux analysis: a universal tool for measuring in vivo metabolic activity*; Springer-Verlag: Berlin, Heidelberg, 2007.

(86)   Fischer, E.; Sauer, U. *Eur J Biochem* **2003**, *270*, 880-891.

(87)   Frick, O.; Wittmann, C. *Microb Cell Fact* **2005**, *4*, 30.

(88)   Marin, S.; Lee, W. N.; Bassilian, S.; Lim, S.; Boros, L. G.; Centelles, J. J.; FernAndez-Novell, J. M.; Guinovart, J. J.; Cascante, M. *Biochem J* **2004**, *381*, 287-294.

(89)   Lee, W. N.; Boros, L. G.; Puigjaner, J.; Bassilian, S.; Lim, S.; Cascante, M. *Am J Physiol* **1998**, *274*, E843-851.

(90)   Toya, Y.; Ishii, N.; Hirasawa, T.; Naba, M.; Hirai, K.; Sugawara, K.; Igarashi, S.; Shimizu, K.; Tomita, M.; Soga, T. *J Chromatogr A* **2007**, *1159*, 134-141.

(91)   Dauner, M.; Sauer, U. *Biotechnol Prog* **2000**, *16*, 642-649.

(92)   Noh, K.; Gronke, K.; Luo, B.; Takors, R.; Oldiges, M.; Wiechert, W. *J Biotechnol* **2007**, *129*, 249-267.

(93)   van Winden, W. A.; van Dam, J. C.; Ras, C.; Kleijn, R. J.; Vinke, J. L.; van Gulik, W. M.; Heijnen, J. J. *FEMS Yeast Res* **2005**, *5*, 559-568.

(94)   Iwatani, S.; Van Dien, S.; Shimbo, K.; Kubota, K.; Kageyama, N.; Iwahata, D.; Miyano, H.; Hirayama, K.; Usuda, Y.; Shimizu, K.; Matsui, K. *J Biotechnol* **2007**, *128*, 93-111.

(95)   Kind, T.; Tolstikov, V.; Fiehn, O.; Weiss, R. H. *Anal Biochem* **2007**, *363*, 185-195.

(96)   van der Werf, M. J.; Overkamp, K. M.; Muilwijk, B.; Coulier, L.; Hankemeier, T. *Anal Biochem* **2007**, *370*, 17-25.

(97)   Ishii, N.; Nakahigashi, K.; Baba, T.; Robert, M.; Soga, T.; Kanai, A.; Hirasawa, T.; Naba, M.; Hirai, K.; Hoque, A.; Ho, P. Y.; Kakazu, Y.; Sugawara, K.; Igarashi, S.; Harada, S.; Masuda, T.; Sugiyama, N.; Togashi, T.; Hasegawa, M.; Takai, Y.; Yugi, K.; Arakawa, K.; Iwata, N.; Toya, Y.; Nakayama, Y.; Nishioka, T.; Shimizu, K.; Mori, H.; Tomita, M. *Science* **2007**, *316*, 593-597.

(98)   Bolten, C. J.; Kiefer, P.; Letisse, F.; Portais, J. C.; Wittmann, C. *Anal Chem* **2007**, *79*, 3843-3849.

(99)   Wittmann, C.; Kromer, J. O.; Kiefer, P.; Binz, T.; Heinzle, E. *Anal Biochem* **2004**, *327*, 135-139.

(100)   Mashego, M. R.; van Gulik, W. M.; Vinke, J. L.; Heijnen, J. J. *Biotechnol Bioeng* **2003**, *83*, 395-399.

(101)   Schaub, J.; Schiesling, C.; Reuss, M.; Dauner, M. *Biotechnol Prog* **2006**, *22*, 1434-1442.

(102)   Mashego, M. R.; Rumbold, K.; De Mey, M.; Vandamme, E.; Soetaert, W.; Heijnen, J. J. *Biotechnol Lett* **2007**, *29*, 1-16.

(103)   Kimball, E.; Rabinowitz, J. D. *Anal Biochem* **2006**, *358*, 273-280.

(104)   Maharjan, R. P.; Ferenci, T. *Anal Biochem* **2003**, *313*, 145-154.

(105)   Rabinowitz, J. D.; Kimball, E. *Anal Chem* **2007**, *79*, 6167-6173.

(106)   Ritter, J. B.; Genzel, Y.; Reichl, U. *Anal Biochem* **2008**, *373*, 349-369.

(107)   Katajamaa, M.; Oresic, M. *J Chromatogr A* **2007**, *1158*, 318-328.

(108)   Katajamaa, M.; Miettinen, J.; Oresic, M. *Bioinformatics* **2006**, *22*, 634-636.

(109)   Katajamaa, M.; Oresic, M. *BMC Bioinformatics* **2005**, *6*, 179.

(110)   Broeckling, C. D.; Reddy, I. R.; Duran, A. L.; Zhao, X.; Sumner, L. W. *Anal Chem* **2006**, *78*, 4334-4341.

(111)   Arita, M. *Brief Funct Genomic Proteomic* **2004**, *3*, 84-93.

(112)   Mendes, P. *Brief Bioinform* **2002**, *3*, 134-145.

(113)   Styczynski, M. P.; Moxley, J. F.; Tong, L. V.; Walther, J. L.; Jensen, K. L.; Stephanopoulos, G. N. *Anal Chem* **2007**, *79*, 966-973.

(114)   Fischer, K.; Hoffmann, P.; Voelkl, S.; Meidenbauer, N.; Ammer, J.; Edinger, M.; Gottfried, E.; Schwarz, S.; Rothe, G.; Hoves, S.; Renner, K.; Timischl, B.; Mackensen, A.; Kunz-Schughart, L.; Andreesen, R.; Krause, S. W.; Kreutz, M. *Blood* **2007**, *109*, 3812-3819.

(115)   Birkemeyer, C.; Luedemann, A.; Wagner, C.; Erban, A.; Kopka, J. *Trends Biotechnol* **2005**, *23*, 28-33.

(116)   Mashego, M. R.; Wu, L.; Van Dam, J. C.; Ras, C.; Vinke, J. L.; Van Winden, W. A.; Van Gulik, W. M.; Heijnen, J. J. *Biotechnol Bioeng* **2004**, *85*, 620-628.

(117)   Nasution, U.; van Gulik, W. M.; Kleijn, R. J.; van Winden, W. A.; Proell, A.; Heijnen, J. J. *Biotechnol Bioeng* **2006**, *94*, 159-166.

(118)   Ohnesorge, J.; Neususs, C.; Watzig, H. *Electrophoresis* **2005**, *26*, 3973-3987.

(119)   Ohnesorge, J.; Sanger-van de Griend, C.; Watzig, H. *Electrophoresis* **2005**, *26*, 2360-2375.

(120)   Soga, T.; Heiger, D. N. *Anal Chem* **2000**, *72*, 1236-1241.

(121) Soga, T.; Imaizumi, M. *Electrophoresis* **2001**, *22*, 3418-3425.

(122) Soga, T.; Ueno, Y.; Naraoka, H.; Matsuda, K.; Tomita, M.; Nishioka, T. *Anal Chem* **2002**, *74*, 6224-6229.

(123) Hardenborg, E.; Zuberovic, A.; Ullsten, S.; Soderberg, L.; Heldin, E.; Markides, K. E. *J Chromatogr A* **2003**, *1003*, 217-221.

(124) Reijeng, J. C.; Martens, J. H.; Giuliani, A.; Chiari, M. *J Chromatogr B Analyt Technol Biomed Life Sci* **2002**, *770*, 45-51.

(125) Ullsten, S.; Danielsson, R.; Backstrom, D.; Sjoberg, P.; Bergquist, J. *J Chromatogr A* **2006**, *1117*, 87-93.

(126) Center for Drug Evaluation and Research, C. f. V. M., Rockville, Maryland; US Department of Health and Human Services, U., Ed., 2001.

(127) Kaspar, H.; Dettmer, K.; Gronwald, W.; Oefner, P. J. *J Chromatogr B Analyt Technol Biomed Life Sci* **2008**.

(128) Caspi, R.; Foerster, H.; Fulcher, C. A.; Hopkinson, R.; Ingraham, J.; Kaipa, P.; Krummenacker, M.; Paley, S.; Pick, J.; Rhee, S. Y.; Tissier, C.; Zhang, P.; Karp, P. D. *Nucleic Acids Res* **2006**, *34*, D511-516.

(129) Gatenby, R. A.; Gillies, R. J. *Nat Rev Cancer* **2004**, *4*, 891-899.

(130) Douglas, D. B.; Akiyama, Y.; Carraway, H.; Belinsky, S. A.; Esteller, M.; Gabrielson, E.; Weitzman, S.; Williams, T.; Herman, J. G.; Baylin, S. B. *Cancer Res* **2004**, *64*, 1611-1620.

(131) T'Kindt, R.; Storme, M.; Deforce, D.; Van Bocxlaer, J. *J Sep Sci* **2008**.

(132) Wiechert, W. *Metab Eng* **2001**, *3*, 195-206.

(133) Fischer, E.; Zamboni, N.; Sauer, U. *Anal Biochem* **2004**, *325*, 308-316.

(134) Mancuso, A.; Sharfstein, S. T.; Fernandez, E. J.; Clark, D. S.; Blanch, H. W. *Biotechnol Bioeng* **1998**, *57*, 172-186.

(135) Miccheli, A.; Tomassini, A.; Puccetti, C.; Valerio, M.; Peluso, G.; Tuccillo, F.; Calvani, M.; Manetti, C.; Conti, F. *Biochimie* **2006**, *88*, 437-448.

(136) Bak, L. K.; Waagepetersen, H. S.; Melo, T. M.; Schousboe, A.; Sonnewald, U. *Neurochem Res* **2007**, *32*, 671-680.

(137) Timischl, B.; Dettmer, K.; Kaspar, H.; Thieme, M.; Oefner, P. J. *Electrophoresis [in press]* **2008**.

(138) Fong, S. S.; Nanchen, A.; Palsson, B. O.; Sauer, U. *J Biol Chem* **2006**, *281*, 8024-8033.

(139) Grabar, T. B.; Zhou, S.; Shanmugam, K. T.; Yomano, L. P.; Ingram, L. O. *Biotechnol Lett* **2006**, *28*, 1527-1535.

(140) Zhang, X.; Jantama, K.; Moore, J. C.; Shanmugam, K. T.; Ingram, L. O. *Appl Microbiol Biotechnol* **2007**, *77*, 355-366.

(141) Weber, J.; Kayser, A.; Rinas, U. *Microbiology* **2005**, *151*, 707-716.

(142) El-Mansi, M.; Cozzone, A. J.; Shiloach, J.; Eikmanns, B. J. *Curr Opin Microbiol* **2006**, *9*, 173-179.

(143) Nanchen, A.; Schicker, A.; Revelles, O.; Sauer, U. *J Bacteriol* **2008**, *190*, 2323-2330.

(144) Prasad Maharjan, R.; Yu, P. L.; Seeto, S.; Ferenci, T. *Res Microbiol* **2005**, *156*, 178-183.

(145) Bautista, J.; Satrustegui, J.; Machado, A. *FEBS Lett* **1979**, *105*, 333-336.

14 Appendix

Table S1. List of average CE migration times in ascending order, observed masses, database hits, peak area ratios of wt over *PntAB-UdhA* under three different growth conditions, and ability of each of the three analytical methods employed to detect, respectively quantitate a given *E. coli* metabolite.

Average MT[a] [s]	Observed mass	Database ID	LB (Stat)	M9 (Log)	M9 (Stat)	Comment[b]	CE-TOF-MS[c]	GC-(AA)-MS[c]	GC-MS[c]
466.1	129.0196	Itaconate	0.02*	n.d.	0.4*	☑	#		
466.8	115.0045	Fumarate	0.15*	0.46*	0.47*	☑	#		x
468.1	87.00676	Pyruvate	0.31*	n.d.	0.27	☑	#		x
468.7	306.9631		0.12*	0.07*	0.23*	$C_5H_9O_{11}P_2$ / $C_6H_5N_4O_7P_2$	#		
468.9	111.0089	Furoic acid	0.05*	7.8*	0.13*	Formula ok!	#		
469.4	191.0207	Citrate	0.005*	0.01*	0.08*	☑	#		x
467.8	192.0255		0.05*	n.d.	n.d.	Isotope of Citrate	#		
469.5	270.0561		0.34*	n.d.	0.26*	$C_6H_{12}N_3O_9$	#		
472.4	246.9436	Cyclic 2,3-bisphospho-glycerate	0.1*	0.07*	0.26*	Formula ok	#		
472.8	149.01	Tartaric acid	0.64	n.d.	n.d.	☑	#		
472.9	193.0265	Glucuronate	0.04*	0.15*	0.12*	☑	#		
473.8	196.056	3-hydroxytyrosine	n.d.	n.d.	0.98	Isotope of glucuronate	#		
474.4	136.0349	Malate-d3 (IS)	0.81	0.78	1	☑	#		
475.1	135.0282	Hypoxanthine	0.93	1.29	0.95	Isotope of Malate-d3	#		
475.3	133.0162	Malate	0.17*	0.27*	0.38*	☑	#		x
476.4	134.0213	Homocysteine	0.08*	0.63	0.47*	Isotope of Malate	#		
474.7	137.0374	Urocanate,	0.81	n.d.	0.97	Formula ok!	#		
473.8	138.0391		0.91	1.01	0.98	Isotope of urocanate	#		
474.7	405.0313		0.03*	0.15*	0.05*	$C_{13}H_{14}N2O_{11}P$	#		
475.3	173.009	cis-Aconitate	0.01*	0.09*	0.17*	☑	#		x
476.2	145.0131	alpha-Ketoglutarate	0.96	0.58	1.08	☑	#		x
476.5	121.0453	Succinate-d4 (IS)	0.96	0.9	1.06	☑	#		
476.7	123.0496	4-Hydroxybenzyl alcohol	1.04	1.09	0.84	Succ-d4-$^{13}C_2$	#		
476.7	122.0489		0.96	1.08	1.1	Succ-d4-^{13}C	#		

477.4	120.0393		0.93	1.03	0.98	Succ-d3	#	
477.7	117.0197	Succinate	0.71*	0.65*	0.54*	☑	#	x
477.6	118.0225	3-Nitropropanoate	n.d.	n.d.	0.83	Isotope of succinate	#	
479.4	281.062		0.47*	0.92	n.d.	$C_8H_{13}N_2O_9$	#	
479.7	265.0784	p-Aminobenzoyl glutamate	1	n.d.	0.96	Formula ok!	#	x
493.3	147.0314	2-Hydroxyglutarate, citramalate	1.06	0.73	2.06*	Formula ok!	#	
494.7	246.0617	N2-succinylglutamate	0.58*	0.42*	0.3*	Formula ok!	#	
495.4	247.0647	5-Hydroxyindoleacetylglycine	n.d.	n.d.	0.27*	Isotope of N2-succinyl glutamate	#	
496.0	205.0347	2-Methylcitrate, methylisocitrate	0.07*	0.03*	0.26*	Formula ok!	#	
496.6	184.9863	Phosphoglycerate	0.48*	1.87*	0.72*	☑	#	x
497.0	166.9729	Phosphoenolpyruvate	0.58*	0.88	0.72	☑	#	
499.3	131.0379	4-hydroxy-2-ketovalerate, 2-aceto-lactate	0.76	n.d.	n.d.	Formula ok!	#	
499.4	260.0823		1.32	0.28*	0.15*	$C_9H_{15}N_3O_4P$	#	
508.0	105.0201	Glycerate	0.1*	1.34	0.81	Formula ok!	#	x
512.5	119.0364	Dihydroxybutyric acid, Dihydroxyisobutyric acid	1.06	0.79	0.98	Formula ok!	#	
517.6	175.037	N-carbamoyl-aspartate,	0.21*	132.05*	2.36*	Formula ok!	#	
518.7	176.0384	N-Formyl-L-methionine	n.d.	22.72*		Isotope of N-Carbamoyl aspartate	#	
520.2	174.0407	N-Acetylaspartate , 2-Amino-3-oxoadipate	0.4*	10.55*	2.55*	Formula ok!	#	
530.8	181.0508	3-(2,3-dihydroxyphenyl)propionate	2.93*	n.d.	2.61*	Formula ok!	#	
533.7	188.059	N-acetyl-glutamate	7.96*	1.12	n.d.	Formula ok!	#	
534.7	159.0712	Dimethylglutarate	1.26			Formula ok!	#	
543.5	168.9915	Glyceraldehyde-3-P, dihydroxy-acetone-P	0.35*	n.d.	0.65	Formula ok!	#	
552.2	171.0088	Glycerol-phosphate	0.61*	7.66*	2.28*	☑	#	x
556.1	363.0555		0.93	1.02	0.94		#	
554.9	365.0531	Phosphoribosyl-formamido-carboxamide	1.22	0.76	1.31	Isotope of 363.0555	#	
557.8	364.0588		0.89	0.97	n.d.	Isotope of 363.0555	#	
556.1	301.0558	PIPES (IS)	1	1.04	0.96	☑	#	
555.7	150.0222		0.97	0.98	0.98	PIPES 2-	#	
555.3	150.5244		1.05	0.97	1	Isotope of PIPES 2-	#	
556.7	304.0543		0.81	0.76	n.d.	Isotope of PIPES	#	
557.9	302.058		0.96	1.21	0.96	Isotope of PIPES	#	
557.0	303.053		0.94	1.01	0.94	Isotope of PIPES	#	

556.8	257.0799	3-Methyluridin , ribosylimidazoleacetic acid	1.81*	0.51*	0.85	Formula ok!	#	
557.3	305.0484		2.17*	n.d.	n.d.	$C_7H_9N_6O_8$	#	
557.4	91.03743	Glycerol	0.47*	1	0.97	Formula ok!	#	
557.4	412.8818		n.d.	3.83*	7.63*	$C_9H_4O_{13}P_3$	#	x
559.0	110.9865	Hydroxymethylphosphonate	n.d.	1.48	1.13	Formula ok!	#	
568.8	229.0125	Pentose-P	0.54	n.d.	0.73	☑	#[d]	
569.4	255.2328	Palmitate	0.74	1.18	n.d.	Formula ok!	#	x
570.3	92.03308	Lactate-$^{13}C_3$ (IS)	1.04	1	1	☑	#	
570.9	89.02238	Lactate	1.27	4.07*	1.31*	☑	#	x
583.4	403.000	UDP	0.29*	n.d.	1.48	☑	#	
586.7	402.0107	CDP	0.28	1.23	1.95*	☑	#	
587.6	115.0423	2-Keto-isovalerate	1.02	1.02	1.2	Formula ok!	#	
588.0	579.0282	UDP-D-glucuronate	0.56	0.97	0.91	☑	#	
594.1	390.9012		0*	6.82*	23.87*	$C_7H_6O_{13}P_3$	#	
594.8	87.0424	2-Methylpropanoate, butyrate, acetoin	2.14	n.d.	1.29	Formula ok!	#	
595.3	122.0303	Picolinate, nicotinic acid	1.98	1.46	0.81	Formula ok!	#	
601.4	112.9865	Acetylenedicarboxylate,	n.d.	n.d.	0.08*	Formula ok!	#	
602.7	259.0221	Hexose-P	0.5*	0.35*	1.37*	☑	#[d]	x
600.4	261.0366	Mannitol-1-P, galactitol-1-P, D-sorbitol-6-P	0.34*	0.71	1.38	Isotope of Hex-P	#	
603.4	260.027	O-phospho-L-tyrosine	0.36*	0.37*		Isotope of Hex-P	#	
602.8	182.0154	L-homocysteic acid	4.98*	n.d.	n.d.	Formula ok!	#	
603.2	155.005	Orotate	0.26*	11.65*	1.03	☑	#	x
603.7	156.0096		n.d.	9.05*	n.d.	Isotope of Orotate	#	
603.7	176.9352	Diphosphate,	n.d.	4.27*	3.33	Formula ok!	#	
604.6	129.0566	2-Keto-4-methyl-pentanoate, 2-keto-3-methyl-valerate	0.87	0.74*	0.56*	Formula ok!	#	
605.1	678.0978	UDP-N-acetylmuramate	0.41*	0.55*	0.54*	Formula ok!	#	
605.4	323.0324	UMP	0.36*	1.4	0.98	☑	#	
605.9	426.0241	ADP	0.81	1	1.96*	☑	#	
608.8	287.0541	Orotidine	8.53*	n.d.	1.29	Formula ok!	#	
610.2	289.0331	Sedoheptulose-7-P	0.57*	0.99	1.05	Formula ok!	#	
610.9	131.0364	Glutarate	1.25	0.65*	2.19*	☑	#	
611.1	128.0358	5-Oxo-D-proline	5.95*	1.85*	2.15*	Formula ok!	#	
612.4	164.0364	Methionine sulfoxide	2.99	n.d.	n.d.	Formula ok!	#	

612.5	322.0466	CMP	0.49*	0.67	1.03	☑	#		
613.2	321.0493	dTMP	0.26*	0.77	0.69	☑	#		
616.1	132.032	Aspartate	2.04*	0.83	1.39*	☑	#	x	
616.1	138.9728	Phosphonoacetate, acetylphosphate	n.d.	1.97	1.15	Formula ok!	#		
619.8	175.0621	Isopropylmalate	n.d.	n.d.	24.67	Formula ok!	#		
621.8	346.0586	AMP	0.46*	1.53	1.16	☑	#		x
629.6	236.0965		3.01*	n.d.	0.77	$C_{12}H_{14}NO_4$	#		
632.3	115.077	Caproic acid	1.39*	1.49*	1.1	☑	#		x
633.7	133.0454	2,3-Dihydroxy-isovalerate,deoxyribose, 1-deoxy-xylulose	3.43*	n.d.	1.43	Formula ok!	#		
633.8	228.0544	5-(L-Alanin-3-yl)-2-hydroxy-cis,cis-muconate 6-semialdehyde	3.35*	n.d.	2.66*	Formula ok	#		
634.9	102.0542	Aminobutyrate,	2.75*	1.45*	2.04*	Formula ok!	#		
635.1	258.0401	Glucosamine-6-phosphate	0.29*	1.65	1.63*	☑	#		
635.6	320.1146	Tripeptide	1.42*	1.17	1.37*	Formula ok!	#		
635.7	173.0608	N-formimino-glutamate,	0.37*	1.06	0.55*	Formula ok!	#		
636.1	298.134		2.38*	1.14	n.d.	$C_{14}H_{20}NO_6$	#		
635.9	151.0779	Glutamate-d5 (IS)	0.91	1.01	0.97	☑	#		
636.3	146.0475	Glutamate	2.78*	2.11*	2.33*	☑	#	x	
635.4	147.0538	Dihydrocoumarin	2.7*	1.32	2.16*	Isotope of Glutamate	#		
637.0	148.0501	Methionine, penicillamine	7.1*	1.1	2.16*	Isotope of Glutamate	#		
636.3	168.0291	2,3-Dihydrodipicolinate,	1.94*	1.3	1.57*	Formula ok!	#		
637.8	121.0312	Benzoate	2.74*	2.17*	1	☑	#		x
638.4	315.0839		4.5*	1.97*	2.68*	$C_{12}H_{15}N_2O_8$	#		
638.7	316.0879	Desmethylnizatidine	9.95*	n.d.	2.74*	Isotope of 315.0839	#		
636.3	595.6624		0.69	1.48	1.26	1193.32 2-	#		
634.6	596.1637		0.6*	1.25	1.27	Isotope of 595.67	#		
640.5	596.6722		0.63*	1.68	1.63	Isotope of 595.67	#		
640.5	597.1724		0.87	n.d.	n.d.	Isotope of 595.67	#		
641.2	349.0273	N-Phospho-D-lombricine	15.36*	n.d.	2.69*	Formula ok!	#		
643.2	742.0748	NADP/NADPH	1.52	1.62	5.01*	☑	#		
645.3	129.0915	Heptanoic acid	1.51	1.51*	n.d.	☑	#		
645.3	154.0123		1.77*	0.79	1.13	$C_6H_4NO_4$	#		
646.9	154.0895		1.13	1.02	0.96	Cinnamic acid–d7 (IS)	#		

644.1	155.0939		1.06	0.94	0.96	Isotope of Cin-d7	#		
642.0	156.096		1.03	0.99	0.97	Isotope of Cin-d7	#		
648.4	282.022		1.08	1.53*	2.29*	$C_6H_8N_3O_{10}$	#		
652.1	565.049	UDP-glucose	0.84*	1.17	2.43*	☑	#		
656.6	375.1053		0.92	1.35*	1	$C_{13}H_{20}N_4O_7P$	#		
657.9	547.0748	dTDP-alpha-L-rhamnose	1.58*	1.32	1.35*	Formula ok!	#		
657.7	548.0807		2.73*	1.83*	1.35	Isotope of dTDP-rhamnose	#		
660.1	143.1081	Octanoate	1.62*	0.83	1.03	☑	#		x
661.4	611.1457	Glutathione disulfide (GSSG)	1.3	1.14	1.11	Formula ok!	#		
660.5	612.1484		1.27	1.54	1.08	Isotope of GSSG	#		
661.2	613.1473		1.54	2.13	n.d.	Isotope of GSSG	#		
662.0	606.0758	UDP-N-acetyl-galactosamine, UDP-N-acetyl-glucosamine, UDP-N-acetyl-mannosamine	0.46*	0.95	0.88	Formula ok!	#		
665.0	302.5349		0.24*	0.93	0.76*	z=2 of 606.0758	#		
665.1	607.0849		0.33*	n.d.	n.d.	Isotope of 606.0758	#		
662.8	101.0592	Valeric acid	n.d.	1.59	1.22	☑	#		x
663.5	333.0295		2.06*	1.96*	0.93	$C_{12}H_{15}O_7P_2$	#		
664.7	334.0289	Nicotinic acid mononucleotide	2.7*	1.84	0.64	Isotope of 333.0295	#		
665.0	145.0609	Glutamine	n.d.	0.35*	1.7*	☑	x	#	
672.0	157.1244	Nonanoic acid	0.99	1.28	1.05	☑	#		
681.7	241.0136	Inositol cyclic phosphate	0.05*	n.d.	1.09	Formula ok!	#		
682.2	225.0902	Porphobilinogen	2.32	n.d.	n.d.	Formula ok!	#		
682.8	171.1377	Capric acid	1.39	1.12	0.59	☑	#		x
684.9	328.0623		1.69*	1.29	1.3	$C_{13}H_{15}NO_7P$	#		
686.1	613.1512		3.01*	2.41*	1.98*	$C_{21}H_{35}N_4O_{13}P_2$ / $C_{22}H_{34}N_2O_{16}P$	#		
686.2	306.075	Glutathione	1.89*	1.64	1.44*	☑	#		
686.4	307.0819		1.39	1.47	1.28	Isotope of glutathione	#		
686.5	179.0545	Glucose	0.86*	n.d.	0.36*	☑	x		#
687.2	334.1247	N4-(b-N-Acetyl-glucosaminyl)-asparagine	n.d.	0.53*	0.58*	Formula ok!	#		
698.1	215.0492	Bisnorbiotin	n.d.	1.02	0.95	Formula ok!	#		
704.4	189.0781	3-Hydroxysuberic acid	n.d.	n.d.	6.59*	Formula ok	#		
705.3	218.1047	Pantothenate	1.47*	0.68	1.18	☑	#		x

706.0	101.0222	Succinate semialdehyde, acetoacetate, 2-oxobutanoate	n.d.	n.d.	2.09	Formula ok!	#		
706.8	332.1467	Tripeptide	n.d.	0.58*	0.75*	Formula ok!	#		
709.9	274.1077	Tripeptide	1.43*	0.85	n.d.	Formula ok!	#		
712.1	182.0437	Methyl(2-furoylamino)acetic acid	n.d.	0.24*	0.13*	Formula ok!	#		
717.4	253.0816	Hydroxyfelbamate	1.82*	n.d.	0.29*	Formula ok	#		
717.6	288.1225	Tripeptide	1.63*	1.1	n.d.	Formula ok!	#		
719.8	289.1218	L-Arginino-succinate,	0.88	1.38	0.75	Isotope of Tripeptide (288.12)	#		
719.3	231.098	N2-Succinyl-ornithine	1.4*	0.76	0.45*	Formula ok!	#		
721.5	232.105		1.3	n.d.	0.33*	Isotope of N2-succinyl-ornithine	#		
719.8	275.0936	Dipeptide (Glu-Glu)	0.94	1.37*	1.06	Formula ok!	#		
720.6	88.03704	Alanine	0.51*	n.d.	1.9*	☑	x	#	x
729.9	245.1176	Dipeptide	1.78*	0.26*	0.15*	Formula ok!	#		
728.7	246.1167		n.d.	0.4*	0.13*	Isotope of 245.11	#		
736.5	273.1213	N2-Succinyl-arginine,	2.02*	0.56	0.95	Formula ok!	#		
734.3	274.1271		1.52	n.d.	n.d.	Isotope of N2-succinyl-arginine	#		
740.4	275.1282		5.02*	n.d.	n.d.	Isotope of N2-succinyl-arginine	#		
765.0	328.0532	cyclic-AMP	n.d.	2.69*	n.d.	☑	#		
773.0	159.0788	Alanyl-alanine	1.64	n.d.	n.d.	Formula ok!	#		
777.1	205.0831	Dipeptide (Thr-Ser)	n.d.	0.67	0.74	Formula ok!	#		
799.1	421.0746	Lactose 6-P, trehalose 6-P, maltose 6'-P, sucrose 6-P	n.d.	0.08*	n.d.	Formula ok!	#		
800.2	662.1055	NAD/NADH	0.83	0.65*	1.08	☑	#		
807.5	663.1085	NAAD	1.13	0.79	1.05	Isotope of NADH	#		
809.6	203.1036	Dipeptide (Val-Ser)	n.d.	0.7	0.72	Formula ok!	#		
834.9	217.1188	Dipeptide , N2-(D-1-Carboxyethyl)-lysine	n.d.	0.45*	0.95	Formula ok!	#		
851.1	185.0793		0.92	0.47*	0.96	$C_8H_{13}N_2OS$	#		
854.8	116.0746	Betaine	0.07	n.d.	n.d.	Formula ok!	#		
855.1	130.0886	Leucine	1.5*	1.42	0.17*	☑	x[e]	#	x
855.1	130.0886	Iso-leucine	1.67*	1.26	wt only*	☑	x[e]	#	x
881.6	103.0399	3-Hydroxy-isobutyrate	1.01	n.d.	n.d.	Formula ok!	#		

876.4	104.0431	Serine	0.99	n.d.	n.d.	Isotope of Hydroxy-isobutyrate		#	
882.8	105.0415	Benzaldehyde	0.99	n.d.	n.d.	Isotope of Hydroxy-isobutyrate		#	
881.6	304.1511	(6S)-Hydroxyhyoscyamine	n.d.	0.48*	0.64*	Formula ok!		#	
885.7	319.127		n.d.	0.4*	0.35*	$C_{11}H_{19}N_4O_7$		#	
1060.9	242.0789	Cytidine	n.d.	5.45*	n.d.	☑		#	
1061.4	135.0314	Hypoxanthine	n.d.	n.d.	0.46*	☑		#	
1088.1	243.0618	Uridine	2.07	9.61*	0.58*	☑		#	
1099.5	266.0889	Adenosine	n.d.	7.94*	0.81	☑		#	
1113.5	266.0763	S-Ribosyl-L-homocysteine	n.d.	n.d.	2.11	Formula ok		#	
1115.1	326.1116	Tripeptide	n.d.	11.53*	0.46*	Formula ok!		#	
1145.5	134.048	Adenine	n.d.	1.64	0.46*	☑		#	
1146.4	253.2166	Palmitoleic acid	n.d.	2.47*	10.22*	Formula ok!		#	
1160.9	164.0709	Phenylalanine	1.34	1.36	0.49*	☑	x	#	
1163.1	155.0004	2,5-Furandicarboxylic acid	n.d.	3.22	0.81	Formula ok		#	
1186.6	145.0977	Lysine	1.22*	1.01	0.29*	☑	x	#	
1190.1	154.0623	Histidine	1.12	0.67*	n.d.	☑	x	#	
1192.6	111.0187	Uracil	n.d.	n.d.	0.49*	☑		#	
1194.7	178.0715	Glucosamine	n.d.	0.45*	0.58	Formula ok!		#	
1206.5	118.0478	Threonine	n.d.	0.9	2.97*	☑	x	#	
1207.8	167.0644	5-(3-Methyltriazen-1-yl)imidazole-4-carboxamide	n.d.	1.21	0.82	Formula ok!		#	
1208.2	121.0377	Nicotinamide	n.d.	0.63	0.97	☑		#	x
1211.4	176.0921		n.d.	0.49*	0.68	$C_8H_{10}N_5$		#	
1223.9	135.0645	5,6,7,8-Tetrahydropteridine	n.d.	n.d.	1.22	Formula ok!		#	
1231.4	114.0567	Proline	0.16*	0.82	1.49*	☑	x	#	
1234.9	341.1123	Sucrose	1.82*	0*	0*	☑	x[f]		#
1234.9	341.1123	Maltose	PU only*	0*	0*	☑	x[f]		#
1237.5	116.0724	Valine	1.46*	1.33	1.78	☑	x	#	x
1237.5	173.1055	Arginine	n.d.	1.58	1.34	☑		#	
1244.6	123.0484	4-Hydroxybenzyl alcohol	n.d.	n.d.	1.01	☑		#	
1248.2	435.2508	1-oleyl-2-lyso-phosphatidate	n.d.	n.d.	1.17	Formula ok!		#	
1255.6	286.0412	Indole-3-glycerol-phosphate	n.d.	1.03	0.94	Formula ok		#	

1318.5	405.2667	3,12-dihydroxy-7-oxo-5-beta-cholanate	n.d.	n.d.	0.87	Formula ok	#		
1323.4	89.01938	Glyceraldehyde, dihydroxy-acetone	n.d.	n.d.	1.18	Formula ok!	#		
1332.5	227.2012	Myristic acid	n.d.	2.05	1.01	Formula ok!	#		x
1339.0	160.0636	Aminoadipic acid	4.9*	0.12*	0.09*	☑	x	#	
1347.9	181.0088	1,3-Glyceryl dinitrate	n.d.	0.97	1.28	Formula ok	#		
1354.1	151.0608	Arabitol, ribitol	1.21	0.97	1.07	Formula ok!	#		

* Significant difference between strains (p<0.05 in Student's t-test)
a) MT: corrected migration time in seconds
b) ☑ - Consistent with commercial standard: Formula ok – generated formula consistent with database hit
c) # - Method that yielded the quantitative data; x – alternative methods confirming metabolite identity
d) Not resolved in MZMine
e Leucine and iso-leucine migrated as a single peak in CE
f) Maltose and sucrose migrated as a single peak in CE

15 Curriculum Vitae

Personal Data:

Name: Birgit Timischl

Date of birth: 13.02.1979

Nationality: Austria

Education:

08/2005 – present:	Natural Sciences Graduate Student at the Institute of Functional Genomics, University of Regensburg, Germany (Advisor: Prof. Peter Oefner)
04/2004 – 07/2005:	Visiting researcher at the Stanford Genome Technology Center, Stanford University, USA
10/2003:	Dipl. Ing. (Masters degree) in Technical Chemistry (with distinction)
09/2002 – 03/2003:	Diploma thesis at the University of Fribourg, CH, on 'The role of sphingolipids in surface stabilization of the yeast plasma membrane H^+-ATPase' (Advisor: Prof. Roger Schneiter)
01/2002 – 06/2002:	Exchange semester at Lund University, Sweden
1997 – 2003:	Technical chemistry studies at the Technical University of Graz; specialization on biochemistry, biotechnology and food chemistry
1989 – 1997:	High school in Fürstenfeld
1985 – 1989:	Primary school in Fürstenfeld

Stipends and Awards:

2005 – 2007	DOC-FFORTE Scholarship, doctoral scholarship for women in research and technology, Austrian Academy of Science
2004 – 2005	Mobility scholarship of the Austrian Genome Research Program GEN-AU
2002 – 2003	Scholarship „Kurzfristige wissenschaftliche Arbeiten im Ausland", Austrian Federal Ministry for Education, Science and Culture
2001	Erasmus scholarship for 6 months at Lund University, Sweden

16 Publications and Presentations

16.1 Publications

[1] **Timischl B**, Dettmer K, Oefner P.
A simple and rapid method for the analysis of glycolytic flux distribution in *Escherichia coli.*
Analytical Chemistry [in preparation]

[2] [*]Dietl K, [*]Renner K, [*]**Timischl B**, Dettmer K, Dorn C, Hellerbrand C, Oefner PJ, Andreesen R, Kreutz M, Gottfried E.
Lactic acid inhibits TNF-alpha secretion of human monocytes via suppression of glucose metabolism.
J Immunol [submitted]

[3] **Timischl B**, Dettmer K, Oefner P.
Hyphenated mass spectrometry in the comprehensive analysis of the central carbon metabolism.
Analytical and Bioanalytical Chemistry **391,** 895-898 (2008)

[4] **Timischl B**, Dettmer K, Kaspar H, Thieme M, Oefner P.
Development of a quantitative, validated CE-TOF-MS method with integrated high-confidence analyte identification for metabolomics.
Electrophoresis **29,** 2203-2214 (2008)

[5] Fischer K, Hoffmann P, Voelkl S, Meidenbauer N, Ammer J, Edinger M, Gottfried E, Schwarz S, Rothe G, Hoves S, Renner K, **Timischl B**, Mackensen A, Kunz-Schughart L, Andreesen R, Krause S., Kreuz M.
Inhibitory effect of tumor cell derived lactic acid on human T cells.
Blood **109**(9), 3812-3819 (2007).

[6] **Timischl B**, Dettmer K, Oefner P.
Metabolomics – An integral component of functional genomics.
In: Fuerst W & Bauernschmitt J. (Hrsg). *Biotechnologie in Bayern*, 26-28 (2007).

[7] Gaigg B, **Timischl B**, Corbino L, and Schneiter R.
Synthesis of sphingolipids with very long-chain fatty acids but not ergosterol is required for routing of newly synthesized plasma membrane ATPase to the cell surface of yeast.
J Biol Chem **280**(23), 22515-22 (2005).

16.2 Oral presentations

CE-TOF-MS based quantitative metabolomics.
    18. Doktorandenseminar des AK Separation Science, Hohenroda (2008)
    Award for best presentation (3rd place)

Quantitative analysis of the central carbon metabolism by capillary electrophoresis – mass spectrometry.
    Baygene Meeting, Regensburg (2007)

Method development for metabolic profiling.
    Invited talk at the Institute of Genomics and Bioinformatics, TU Graz (2006)

17 Summary

The combination of several different approaches is required for the comprehensive analysis of all the small molecules in a biological system, termed metabolomics. Here, we present two different strategies towards this final goal. A capillary electrophoresis - mass spectrometry method was developed and validated for the quantitative analysis of negatively charged metabolites by making use of the high mass accuracy and the quantitation capabilities of a time-of-flight mass analyzer in combination with automated feature extraction and database search. The method was used to elucidate metabolic changes in the *Escherichia coli* deletion mutant *UdhA-PntAB* that lacks nicotinamide nucleotide transhydrogenase function, under both stationary and exponential growth conditions. Concomitant analyses with two different gas chromatography-mass spectrometry methods allowed not only cross-validation of the quantitative results obtained by the various methods, but also led to a more comprehensive coverage of the *E. coli* metabolome.

In the second approach we introduced rapid resolution ion pair-liquid chromatography - tandem mass spectrometry for the analysis of the isotopomer distribution in intermediates of the upper and the lower part of the central carbon metabolism, including glycolysis, PPP, TCA, and related pathways, in cells grown in culture media supplemented with $1,2\text{-}^{13}C_2$-glucose. Additionally, reliable quantification of metabolites was achieved by the use of stable-isotope labeled internal standards. The applicability of the method was tested by comparing isotopomer distributions in central carbon intermediates in the two *Escherichia coli* strains described above. We observed in the mutant strain the expected increase in carbon flux through the oxidative PPP in concordance with published gas chromatography – mass spectrometry data on the mass isotope distribution in proteinogenic amino acids. We also observed in the mutant a relative increase in carbon flux through the NADPH-generating isocitrate dehydrogenase reaction of the TCA cycle, while the carbon flux through the glyoxylate pathway was decreased in comparison to

wild type *E. coli*. Further, absolute levels of TCA cycle intermediates were generally higher in the mutant than the wild type. Finally, the mutant also exhibited a decreased flux through the NADPH consuming methylglyoxal pathway, an offshoot of glycolysis, which was inferred indirectly from an imbalance of unlabeled and doubly labeled extracellular lactate. The data support the utility of RR-IP-LC-MS/MS in the rapid and simple screening for alterations in flux distributions in the central carbon metabolism. It does not require any prior information on the biological system of interest and, therefore, can be applied to any prokaryotic or eukaryotic cell.

18 Zusammenfassung

Die umfassende Analyse aller kleinen Moleküle eines biologischen Systems, genannt Metabolomics, kann derzeit nur durch die Kombination von mehreren unterschiedlichen analytischen Methoden erfolgen. In der vorliegenden Arbeit werden zwei Strategien in diesem Kontext vorgestellt. Eine Methode basierend auf Kapillarelektrophorese – Massenspektrometrie wurde für die quantitative Analyse von anionischen Metaboliten entwickelt und validiert. Als besonders vorteilhaft stellte sich die Verwendung eines hochauflösenden Flugzeitmassenspektrometers dar, wodurch die robuste Quantifizierung von ausgewählten bekannten Metaboliten mit automatisierter Detektion von zusätzlichen potentiellen Metaboliten und anschließender Datenbanksuche verbunden werden konnte. Die Methode wurde zur Untersuchung von metabolischen Veränderungen unter verschiedenen Wachstumsbedingungen in der *Escherichia coli* Deletionsmutante *UdhA-PntAB*, in welcher beide Nikotinamidnukleotid-Transhydrogenasen fehlen, eingesetzt. Durch die gleichzeitige Analyse der Proben mit zwei verschiedenen Gaschromatographie – Massenspektrometrie basierten Methoden, konnte zum einen eine Vergleichsprüfung der quantitativen Ergebnisse erfolgen, zum anderen konnte auch eine umfassendere Abdeckung des *E. coli* Metaboloms erzielt werden.

Des weiteren wurde eine Methode basierend auf Ionenpaar – Flüssigkeitschromatographie gekoppelt mit einem Tandemmassenspektrometer entwickelt und für die Analyse der Isotopomerenverteilung in Intermediaten des gesamten zentralen Kohlenstoffstoffwechsels, i.e. Glykolyse, PPP, TCA und davon abzweigende Stoffwechselwege, eingesetzt. Dafür wurden Zellen verwendet, welche in Nährmedien inkubiert wurden, die mit $[1,2-^{13}C_2]$Glukose supplementiert worden waren. Zusätzlich konnten mit dieser Methode ausgewählte Metaboliten mit Hilfe von isotopenmarkierten internen Standards zuverlässig und genau quantifiziert werden. Die Methode wurde auf ihre Anwendbarkeit getestet, indem die Isotopomerenverteilung innerhalb des zentralen Kohlenstoffstoffwechsels in den beiden oben genannten *E. coli*

Stämmen gemessen wurde. In der Mutante konnte eine Verschiebung der Stoffwechselflüsse in Richtung des oxidativen PPP beobachtet werden, was sehr gut mit publizierten GC-MS Daten korrelierte. Außerdem wurde in der Mutante ein relativer Anstieg des Flusses durch den Isozitrat-Dehydrogenase katalysierten Teil des Zitratzyklus beobachtet, im Zuge dessen auch NADPH generiert wird, während der Fluss durch den Glyoxylatweg im Vergleich zum Wildtyp verringert war. Des weiteren wurden in der Mutante zumeist höhere Konzentrationen an Zitratzyklus-Intermediaten gefunden. Schließlich zeigte die Mutante zudem einen verringerten Fluss durch den Methylglyoxalweg, in welchem NADPH verbraucht wird und welcher einen Seitenweg zur Glykolyse darstellt. Diese Daten unterlegen die Eignung der entwickelten Methode als schnelle und einfache Strategie für das Screening von Veränderungen in den Flüssen des zentralen Kohlenstoffstoffwechsels. Man benötigt dafür keinerlei vorausgehende Informationen über das untersuchte biologische System, weshalb die Methode auf jede prokaryotische oder eukaryotische Zelle angewendet werden kann.