Discussion and Summary
General discussion

The aim of this thesis project was to develop different aspects of a metabolomics workflow to study metabolism in different biological systems. I provide my personal evaluation and future perspectives per chapter.

**Quenching of metabolism by cold mixtures of organic acids**

The presented quenching method is the first organic solvent-based method that enables quantitative analysis of numerous intracellular metabolites in L. lactis grown on chemically defined media with glucose. The described approach can be applicable to the development of a new quenching method for other organisms.

There is a general belief that membrane integrity has to be preserved to successfully separate intracellular and extracellular fractions. We showed that a compromised membrane might not be as big a problem as anticipated. Because intracellular metabolites—especially phosphorylated metabolites—show limited solubility in the quenching solution, they precipitate and due to the centrifugation step end up in the cellular fraction.

Based on our results I suggest that further development of new quenching approaches for analysis of intracellular metabolites should rather focus on limiting solubility of metabolites instead of preventing membrane disintegration. Indeed, we tested membrane integrity of yeast after quenching with methanol—the standard and validated quenching method—and our data suggest a compromised yeast membrane. There is one caveat, though, as we could only run the flow cytometry at room temperature, and so we cannot exclude that the membrane was only compromised after warming the cell up again. Further detailed studies in yeast should be performed to sort this out.

I also suggest that the reason why quenching methods are not applicable to other organisms could be caused by the difference in growth media composition. If precipitation of metabolites plays a dominant role—as we suggest—then media composition will have a great impact on the efficacy of the quenching method, as metabolite solubility depends rather unpredictably on medium composition. I therefore guess that the organic solvent mix developed for quenching of L. lactis is not L. lactis specific, but specific for the growth medium (CDMPC with glucose) that was used for L. lactis. I suggest that yeast and E. coli grown in the same media could be quenched as described for L. lactis, a hypothesis that remains to be tested.
Apart from the medium composition, during quenching the temperature plays a pivotal role: not only does an increase of temperature enhance solubility of metabolites, it may also allow enzymatic activity. Detected metabolite concentration therefore would not represent metabolism at the time of sample collections. I preformed quenching at -40 °C in glass vials and acetonitrile dry ice mixture. However, laboratory equipment such as a tabletop centrifuge do not cool down to those temperatures. To compensate for this, the centrifugation rotor needed to be cooled down externally, and this is time consuming. I also noticed that the rotor temperature rises quickly during the centrifugation step and therefore limits the number of processed samples. One way to ensure temperatures below zero during the quenching procedure could be to perform entire experiment in a room cooled to -20 °C. An alternative would be to use fast filtration instead of centrifugation. But even quenching with a fast filtration step is time and manpower demanding. From our data we can see that keeping an intact membrane during quenching is difficult and therefore a separation of the intracellular and extracellular fraction without leakage is impossible, unless we would use cryoprotectants, potentially. But since those compounds, such as glycerol, negatively interfere with analysis and sample preparation, they are hardly used. Major reason for intracellular and extracellular fraction separation is sample complexity reduction. But some technological developments might be already capable to tackle this problem such as an ion mobility-MS (IM-MS). Combined with LC, this technology provides, next to chromatographic and mass spectrometric separation, ion gas-phase separation based on ion mobility in an electric field and potentially enables separation of isobars, isomers, conformers and enantiomers.

So, do we still need quenching in cold organic solvents, if we cannot really separate intracellular and extracellular metabolites? Indeed, novel high throughput approaches analyze culture samples directly with MS without any sample treatment. These approaches are not perfectly quantitative because of uncontrolled matrix effects but allow high throughput suitable for time series analysis or screening. Another example of direct online measurement of metabolites is the development of the intelligent knife (iKnife) that enables tissue real-time mass spectrometric analysis during surgery. But as long as real-time analysis is not accessible for all kinds of situations, we have to be able to quench metabolism at the moment of sampling, to prevent any enzymatic or catalytic activity. Therefore I believe our insights into the quenching by cold organic solvents, is going to be useful irrespective of the analytical developments.
Setting up and comparison of different platforms for amino acid analysis

In this chapter we compared three analytical platforms used for quantitative analysis of amino acids actively used in our laboratory. For each platform we used different separation mechanism and different MS instruments. The first platform was based on HILIC separation (Phenomenex Luna-NH2 column (250 × 2.0 mm, 5 µm) with triple quadrupole MS detection (Shimadzu LCMS-8030; 2011); the second platform used reversed phase separation (Agilent Zorbax Eclipse XDB-C18 column; 50 × 4.6 mm, 1.8 µm) in combination with triple quadrupole MS detection (AB-Sciex API 3000; 1998), and the last system utilized gas chromatography separation combined with single quadrupole MS detection (GCMS-QP2010 Plus; 2006).

It is intriguing that both triple quadrupole mass spectrometers performed comparably with respect to the level of sensitivity, even though their release dates differed by 13 years. Despite the Shimadzu instrument’s better scanning speed and faster polarity switching, for the purpose we used it (~50 SRMs, positive ionization mode only with the peak width in tenths of seconds) it performed comparably to the AB-Sciex instrument. The most noticeable difference between these instruments was in the software for instrument control and data processing, which was certainly caused by the age difference between the two. API-Sciex Analyst software (version 1.5.1) was developed in 1998-2000, compared to Shimadzu LabSolution software (version 5.53) that was developed in 2008-2011. Preference of the data acquisition software can be very personal, and it is usually dependent on what was the person’s first software to work with. But data processing and quantitative analysis tools were better evolved in LabSolution. The software is according to me more logical, easier to learn and overall it offers better viewing options.

A great example of software development is Agilent’s change from Empower to Open Lab. I would dare to compare this improvement to changing from a manual to an automated transition in a car. The software has a new format including many new visual features, and it is user friendly. In addition, Agilent invested a lot of time to provide online instruction videos and trainings, which make the software learning curve fast. To give credit to Shimadzu, GC-MS LabSolution software (version 2.71; developed in 1999-2011) including data analysis software is user friendly. It did require software training to get the complete insight in the software options, but it was easy to train other users.
From my experience I would say it is not the highest priority in academia to invest in the latest version of software and that leads to many limitations and delays. I would suggest that during the purchase of any equipment, an agreement or budget should be made available to make sure that for the next 10-15 years the newest version of the software will be used. Also, I hope that vendors will provide better online study materials that help users to develop software knowledge.

**Platform stability and data quality assessment in metabolomics**

In the third chapter we addressed the monitoring and interpretation of analytical platform stability data. During my 10 years of working in six different laboratories, scientific institutes and pharmaceutical companies in the Netherlands, I observed an interesting phenomenon. It appeared relatively common that, during molecular biology experiments, scientist would use positive and negative controls and two sets of size markers, to draw their conclusions from the obtained results. However, once we enter the analytical field, that uses separation techniques combined with mass spectrometric detection, controlling platform stability or assay validity became a matter of personal taste. There are examples where scientists controlled their platform only after cleaning the interface, and others, who may have calibrated mass accuracy, but did not follow signal intensity. Furthermore, some people used quality control samples but did not follow the trends from the obtained data. What was even more intriguing, from my own experience, was that the industrial environment could have suffered an identical problem, but from a different perspective. Due to good manufacturing practice (GMP) and good laboratory practice (GLP) regulations, analytical assays, developed in an industrial environment already had quality controls for assay validity, but setting up acceptance criteria was again a matter of personal preference, e.g. during the quantitative HPLC-UV analysis of viral particles, a control sample, consisting of working standard, is injected during the sample batch and its nominal value is monitored. In conclusion, if an assay is valid, one person could suggest that the average control sample value should be within 85-115% of its expected value, however, another might suggest that all control sample values should be within 85-115% of their expected value.

Further to these observations, and in my opinion, I would like to point out that the analytical field seems to lack clear practical instruction in the setting up of the assay validity criteria. I suggest that this problem needs to be
addressed at different levels. First, good laboratory practice (GLP) and an introduction to good manufacturing practice (GMP) should be a prerequisite at the start of any bachelor degree in science. Students should not be allowed to work in a laboratory till they have received some level of GLP and GMP certification. Second, I believe not only analysis performance but also correct data interpretation, should be part of the analytical education; students should be exposed to and given instruction in more complex data interpretation and statistical analysis. Interpretation of quality controls for the setup of assay validity criteria would be an excellent example. Third, in all scientific fields, more effort should be made towards improving the traceability and control of generated data. I can understand that GMP standards are not necessarily applicable in the academic world, but the implementation of assay protocols, standardized operation protocols, active supervision of student laboratory journals and well-organized data management would be a significant step forward and is likely to be a requirement of future project proposals.

**Transportomics & Alternative approaches for metabolite identification**

In Chapters 4 and 5 we showed how the metabolomics field benefits from a combination with other fields. But what is the true reason for this benefit? Metabolism is still not fully discovered despite 100 years of biochemical studies. One example is the pyruvate transporter into mitochondria that was discovered only recently (4). Reason for this could be that metabolism’s high degree of complexity through the connectivity of metabolic pathways makes it very difficult to draw conclusions from metabolomics data alone. For example ATP participates in more than 180 reactions; NADPH or glutamate are interacting in over 70 reactions (5). What does an increase in these metabolites mean? One way to handle such complex data is to integrate the obtained data into mathematical models through a system biology approach, where they may also be combined with data from other –omics fields. The genome–scale metabolic models (GEMs) contain reconstructions of entire metabolism of a given organism based on genome information and metabolic capabilities of the cell. In addition GEMs have been expanded by transcription and translation (6) and can integrate proteomics data sets (7).

We showed a case where exometabolomics data could be usefully integrated with GEMs, but for intracellular metabolites this is much less clear. The reason is that GEMs use fluxes as variables, not metabolites. For external metabolites, accumulation can be easily translated into a rate, but
to turn intracellular metabolites into fluxes, one would need enzyme kinetics. Such information is not available at genome scale. The only way currently how metabolomics data can be used is through thermodynamic constraints, where the direction of flux is determined through calculation of the (concentration-dependent) Gibbs free energy of the reactions (8).

It is obvious that using GEMs and other modeling approaches is certainly beneficial for biological interpretation and integration of metabolomics data, but I can honestly say that this knowledge and skill set is the domain of system biology departments only. I would suggest that bioinformatics should become a common skill set of any scientist to enable big data sets handling. Another option would be to integrate bioinformaticians into analytical groups to provide this knowledge. Clearly, to move the field forward we need to think and act much more interdisciplinary.
Summary

This thesis describes different parts of the analytical workflow during metabolomics analysis and shows two examples of metabolomics applications. The introduction section explains general terms and provides an overview on why metabolomics studies are performed, how those studies can be subdivided, what analytical platforms can be used and what the general advantages and disadvantages of those platforms are. In addition, this section addresses different steps in sample-processing, analysis and subsequent data interpretation.

The following three chapters provide examples of assay development, sample preprocessing steps and automated data processing, including correction for time related bias. The last two chapters provide an example of the application of metabolomics in order to answer different biological questions.

In Chapter 1 we have investigated quenching, a sample-processing step that provides a fast and complete stop of the cellular metabolism during a metabolomics workflow, necessary when intracellular metabolites are to be measured. We provided in the introduction an overview of quenching approaches and their advantages and disadvantages. In this chapter we tested a set of different organic solvent mixtures as quenching solutions for metabolism analysis of *Lactococcus lactis*. Our goal was to stop *L. lactis* metabolism by cooling it down to -40°C while preserving membrane integrity. From a long list of tested organic solvents, we ended up with two final quenching solutions candidates. To prove that quenching was accomplished without leaking of intracellular metabolites to extracellular fractions, mass balance analysis was used. Our data showed that a 2-methyl-1-propanol based quenching solution is the best choice for an organic-solvent-based quenching solution, that enables quantitative analysis of 28 intracellular and extracellular metabolites. The key to a successful quenching method turned out to be the solubility of compounds in the water and organic phase at low temperatures, not so much the prevention of leakage itself.

Chapter 2 describes the implementation and validation of three analytical platforms for amino acid analysis. In the introduction section of this chapter we provide a long review of analytical platforms used for amino acid analysis describing their advantages and disadvantages. In the experimental section we provide validation data and a comparison of three analytical platforms:
gas chromatography in combination with mass spectrometry, reversed phase chromatography in combination with mass spectrometry and hydrophilic interaction liquid chromatography (HILIC) coupled to mass spectrometry. During the comparison we discuss issues such as a sample pretreatment, platform performance, the use of $^{13}$C-labeled internal standards or the application of methods to a complex biological sample, and describe the overall advantages and disadvantages of all three platforms. We concluded that all three platforms showed comparable results but some of them showed specific advantages.

The effect of instrumental instability, the so-called instrumental drift, on the quality of metabolomics data, was investigated in Chapter 3. To monitor and remove time dependent bias, we developed an R script that enables automated data detrending and correction based on the quality control sample (QC) values observed over time. We showed that when the instrumental drift takes a place it leads to an overall increase in bias in all analytical data including QC, calibration, and biological data. If such biological data were to be used for metabolite flux calculations, they would lead to unnecessary errors or masking of biological effects. We showed that the R-script that was developed could remove such a bias.

We used metabolomics in Chapter 4 to investigate physiological functions of ATP Binding Cassette Subfamily C Member 2 (ABCC2), one member of the ATP-binding cassette (ABC) transporters family. We developed a new approach to discover new substrates of ABC transporters in which urine extracts were incubated with membrane vesicles containing an ABC transporter, after which the compounds transported into the vesicles were analyzed by LC/MS-based metabolomics. Using this transportomics approach we identified many new substrates of ABCC2 such as glucuronides and sulphone-glucuronides of plant-derived xenobiotics. Moreover, we confirmed using vesicular transport experiments that the excretion of these compounds in vivo depends on ABCC2 activity.

An untargeted metabolomics approach was used to investigate the nitrogen metabolism of Bordetella pertussis, the causative agent of Whooping Cough, in Chapter 5. During the metabolomics workflow, we accelerated the time-consuming identification step by combining LC-MS data after scanning the exometabolome of B. pertussis with comprehensive, genome-scale metabolic reconstructions. We reasoned that genome-scale reconstructions provide a context for interpretation and allow for prioritization of candidate metabolites. The idea behind this was that only those metabolites predicted
by the genome-scale model needed to be considered. Using this model-guided filtering lead to 50% success rate in metabolite identification and enabled us to identify eleven new metabolites involved in nitrogen metabolism of *B. pertussis*.

The **general discussion section** of this thesis contains my personal evaluation per chapter and provides suggestions for future developments. Altogether, the methodologies and topics described in this thesis will hopefully help the metabolomics field move forward.

**References**
