1 Introduction
1.1 Metabolomics

Metabolites are small molecules (usually < 1 kDa) that form the molecular fundament of life and are the result of biological systems interacting with their environment. The (poly)carboxylic acid intermediates of the Krebs cycle serve as a good example as they are essential to the cellular respiratory system and regulate energy metabolism, but have lately also found to be bioactive mediators in immunological events. Other examples of metabolites are amino acids, sugars, nucleotides, and lipids. Metabolites can provide a functional readout of the current state of a cellular system, thereby directly visualizing biochemical activity either as biochemical turn-over or in the form of cell signalling molecules, triggering downstream alterations. Metabolomics is defined as ‘the scientific study of the metabolome, or set of metabolites within an organism, cell or tissue’ and in the last decade, metabolomics has emerged as an important field next to the other ‘-omics’ technologies. This can be partly visualized by the number of publications per year on the subject (see Figure 1). The complementary nature of metabolomics to genomics, transcriptomics, and proteomics is a major cause for this development. While alterations in deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and protein concentration as well as other factors such as enzyme occupation or post-translational modification are difficult to analyse in a comprehensive manner, investigating the metabolome is advantageous as it can be seen as the amplification and final result of the aforementioned factors. Also, small changes in particular metabolic fluxes are hard to observe without investigating the metabolome on a molecular level. In summary, it is believed that changes in the genome and/or the proteome in the end will narrow down to alterations of specific metabolites, which possibly are easier to detect as they might reflect the amplification of the
aforementioned up-stream changes. By being able to directly investigate metabolites participating in biochemical processes, disciplines such as systems biology benefit greatly. Also, metabolites might serve as biomarkers of disease progression and prediction⁷, to monitor and diagnose respectively. This can be exemplified by the stable peroxidation products of arachidonic acid, F₂-isoprostanes, being a marker for increased oxidative stress in patients with Alzheimer’s disease⁸-¹⁰.

**Figure 1**
Number of publications found with the keyword 'metabolomics' on ScienceDirect.com in the recent years.
1.2 Analysis in Metabolomics

Basically, two different approaches can be taken when metabolites are to be measured. It is important to determine whether a defined set or as many as possible metabolites are of interest/to be analysed. When earlier research or theory is ground for a hypothesis on what to look for, a targeted approach should be devised. However, when this is not the case or when an unbiased approach is preferred, untargeted analysis might result in a better overview of the metabolic composition of a sample.
1.2.1 Targeted Metabolomics

Although metabolomics is a relatively new term, this manner of performing research is actually the standard for decades\(^5\). Targeted metabolomics is driven by a hypothesis and as the name indicates, only certain molecules are targeted. A daunting amount of publications reviewing strategies for sample preparation and analysis of different metabolites is available. Unfortunately, no universal analytical technique exists that is able to measure all metabolites of interest in a precise and quantitative manner. Therefore, limitations as to what can be measured with appropriate quality have to be taken into account\(^{10-11}\). As only a specific panel of expected endogenous metabolites, related to one or more biological processes, have to be measured for targeted analysis, standards can be used. This enables the quantitative analysis of biological samples and greatly improves specificity. Normally, standards are measured first to optimize separation and detection conditions for the development of the analytical method. Standard curves should then be made (inside and outside the biological matrix) to assess linearity, limits of detection and matrix effects\(^{10}\). In the case that extraction of the metabolites is necessary, the recovery of the extraction has to be determined. An advantage of targeted analysis is the possibility to minimize matrix effects by correcting the resulting peak areas with an internal standard or by dilution\(^1,12\). After validation, samples of different origin, such as healthy and disease or different stages of a disease, can be compared\(^5\).

1.2.2 Untargeted Metabolomics

Untargeted metabolomics is a hypothesis generating approach with the aim of simultaneously analysing as many metabolites as possible in an unbiased manner\(^5\). Detected metabolites during metabolomics studies are usually defined as features which are is typically an \(m/z\) value found in at least three consecutive mass spectra,
aligned with a retention time, having a chromatographic peak shape. Chromatographic separation of complex samples, followed by MS detection is known to be able to analyse most ‘features’, compared to other techniques, such as direct-infusion mass spectrometry (MS) techniques, or nuclear magnetic resonance (NMR)\textsuperscript{5}.\textsuperscript{12} This results from the fact that MS is more sensitive than NMR and in combination with a chromatographic technique, matrix effects such as ionization suppression are reduced. Typically, quadrupole–time-of-flight (Q–TOF) or Orbitrap instruments are used in metabolomics studies, as high mass resolution is recommended. The obtained data of metabolomics studies usually consists of large data files and many software solutions are devised to handle the obtained results. Firstly, peak-alignment algorithms are required to correct for retention time variations, in order to enable comparison between all the measurements. Secondly, the discriminating features found have to be assessed and, preferably, the structure of these compounds must be elucidated. This last objective remains a challenge in many cases\textsuperscript{1}, but metabolite databases such as the human metabolite database (HMDB), containing 41,514 metabolites at the moment\textsuperscript{13}, might provide useful information. Unfortunately, these databases are incomplete for the time being and most features remain unresolved. This is also due to the large differences which can be observed in electrospray ionization (ESI)-MS(MS)spectra between different instruments, as opposed to electron ionization (EI)-MS. EI-MS remains the standard for compound identification by database searches.

### 1.3 Biofluids

Metabolites are found in cells, tissues and biofluids. Biofluids are defined as a biological fluid, which can be excreted, secreted, obtained with a needle or are the result of an ailment, such as urine, saliva, plasma or blister fluid, respectively\textsuperscript{14}.
Analysing metabolites in biofluids remains challenging in many cases and requires a tailor-made approach for each biofluid under investigation, as each matrix (biofluid) comes with its own set of demerits. Often, biological matrices are very complex, as they contain hydrophilic and hydrophobic metabolites, salts, proteins and cells in concentrations spanning many orders of magnitude. Normally, separation and detection instrumentation can be optimized to analyse only limited classes of compounds. Some components will introduce difficulties with the selectivity and/or reproducibility of an analytical method. This renders the direct analysis of all metabolites practically impossible and separation of the metabolites from the matrix as much as possible is advised\textsuperscript{15-17}.

\section*{1.4 Chromatographic techniques}

Separating metabolites from their matrix is possible with chromatography. This a very versatile approach for this purpose based on the difference in partitioning of the metabolites and its matrix between a liquid or gaseous mobile phase and a stationary phase. We can distinguish between liquid chromatography (LC) and gas chromatography (GC).

\subsection*{1.4.1 Liquid Chromatography}

As complex biofluids usually have to be separated in order to reduce sample complexity as well as matrix effects, chromatographic techniques such as LC are employed frequently in small metabolite analysis. For LC, two major separation principles can be utilized: reversed phase LC (RPLC) and hydrophilic interaction LC (HILIC). RPLC is based on the hydrophobic interaction between an analyte and the (non-polar) stationary phase. Typically, an aqueous buffer is the weak solvent and an organic modifier such as methanol or acetonitrile acts as a strong elution solvent.
Chapter 1

This approach has good compatibility with biofluids, as preferably a highly aqueous sample is injected to prevent breakthrough of analytes. RPLC is less suitable for highly polar metabolites, as these are not retained and leave the analytical column in or close to the void volume. This lack of separation renders the analysis less robust and sensitive. On the other hand, HILIC separations are based on the interaction between the analytes and the aqueous layer near the (polar) stationary phase. Different column chemistries such as bare silica, propylamine or propionitrile (cyano) each have their own merits, depending on the compounds of interest. Opposite to RPLC, the aqueous buffer acts as the strong elution solvent, whereas the organic solvent is the weak elution solvent. HILIC is an excellent choice for separation of (very) polar analytes. However, method development for HILIC is less straightforward, as buffers and column chemistries show a strong influence on the obtained results and in addition interact with each other, complicating the situation even more.

Another recent development in the field of metabolomics is the demand for fast run times, as the number of samples per study gradually grows. To still have a sufficient peak capacity, resolution is a point of attention. Next to the employed column chemistry, the most important property of an analytical column is the particle diameter. Much effort is put into the development of small particle sizes with a narrow distribution. State of the art is porous ultra-high performance liquid chromatography (UHPLC) material with 1.7 µm diameter particles, requiring pumps that can cope with the increase in pressure due to these small diameters. An alternative to this UHPLC material is core-shell column material, which has a silica solid core with a thin porous layer. The thin porous layer allows for quick mass transfer kinetics (see Figure 1, the C-term in the Van-Deemter curve) and, therefore, high flow rates, reducing analysis time. Due to the solid core, pressure will not
increase as much as it does by the use of fully porous material. The need for specialized UHPLC pumps can be avoided by using core shell material\textsuperscript{18}.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{van_deemter_curve.png}
\caption{The Van-Deemter curve. The terms describe; $A =$ eddy-diffusion; $B =$ longitudinal diffusion; $C =$ mass transfer coefficient and $v =$ the linear velocity.}
\end{figure}

\subsection{Gas Chromatography}

GC is a very robust alternative to LC. Separation is mainly based on the boiling point of volatiles. However, GC usually requires a more elaborate sample preparation when compared to LC. In most cases, when biofluids are to be analysed, the polar non-volatile analytes are to be extracted and derivatised to ensure GC compatibility. Hyphenation with MS has been a major development enabling highly selective and sensitive targeted as well as untargeted analysis by GC–MS\textsuperscript{19}. One major advantage of GC in combination with electron-ionization mass spectrometry (GC–EI-MS) is the availability and usability of large spectral libraries such as the NIST library containing mass spectra of $>$300,000 components. In combination with the high reproducibility of GC–EI-MS, this can be used for metabolite identification based on characteristic fragmentation\textsuperscript{20}. 

Chapter 1
1.5 **Mass Spectrometry**

Next to electromagnetic spectroscopy techniques such as nuclear magnetic resonance (NMR), MS is the detector of choice in metabolomics. High resolution (HR) instruments, such as a quadruple–time-of-flight (Q–TOF) instrument, and unit-resolution tandem MS instruments, such as a triple quadruple (TQ) instruments, both have become important in the field of metabolomics. While Q–TOF instruments are frequently used for untargeted metabolomics studies, TQ instruments are usually the apparatus of choice in targeted investigations. The hyphenation between chromatographic techniques and MS enables detection of metabolites directly after separation from the matrix. For LC, ionization techniques like electrospray ionization (ESI) and atmospheric-pressure chemical ionization (APCI), and for GC, electron ionization (EI) were great accomplishments, enabling this field. For targeted metabolomics, the sensitive TQ MS instruments are the recommended choice. Selected ion monitoring (SIM), which is also possible in single quadruple MS instruments, or selected reaction monitoring (SRM) render this technique very selective for their targets. Collision induced dissociation (CID) is most used as a fragmentation method for small molecules.²¹

1.6 **Derivatisation**

Unfortunately, not all target analytes are compatible with LC or GC separation and/or with mass spectrometry with EI, ESI or APCI ionization. Derivatisation is the controlled chemical alteration of an analyte, resulting from a reaction with a derivatisation reagent. For GC–MS, derivatisation is a common procedure during sample preparation, to enhance volatility, e.g., by methylation or silylation. However, for LC–MS in combination with soft ionization techniques, derivatisation is preferably avoided, as this renders sample preparation more complicated and laborious and,
therefore, lowers the recovery and reproducibility\textsuperscript{22}. Nevertheless, also in LC, derivatisation can improve analyses by enhancing separation and/or ionization characteristics. Compound classes like aldehydes, ketones and sugars are not or poorly ionized by soft ionization techniques\textsuperscript{23}. Incorporating a functional group such as an amine will strongly enhance the ionization efficiencies for positive ionization mode, whereas a strong acidic functional group will aid the ESI ionization process in negative ionization mode. Additionally, fragmentation properties can be improved by choosing a reagent that selectively fragments at a certain position, enabling the use of e.g. neutral loss scan in tandem MS\textsuperscript{24, 25}. Next to chemical derivatisation for advantageous MS properties, enhancement for chromatography can also be achieved by derivatisation. The partition coefficient of especially polar metabolites might not be readily suitable for separation by generic RPLC. The addition of a phenyl ring will increase the lipophilicity of such hydrophilic compounds and, therefore, significantly improve the chromatographic behaviour\textsuperscript{25}. Finally, when LC is hyphenated to a spectroscopic technique, such as ultra-violet (UV) or fluorescence (FL) detection, a reagent can be chosen to introduce beneficial spectroscopic properties (a chromophore or fluorophore).
1.7 References

Scope

This thesis aims at the development of improved chromatography based methods for the analysis of small metabolites. Several chromatographic techniques are used, with and without derivatisation and in combination with different MS instruments. Focus mainly lies on the targeted analysis of metabolites, such as the carboxylic acids of the Krebs cycle, or malondialdehyde, but also untargeted profiling is touched.

Chapter 1 gives a general overview on the field of metabolomics and its challenges. Different chromatographic and mass spectrometric techniques are briefly introduced within the scope of the analysis of endogenous small molecules. Potentials and limitations of these techniques are shortly discussed, together with derivatisation strategies, as an enhancement of certain technical shortcomings. Also, the two main strategies in metabolic analysis, that is targeted and untargeted metabolomics, are explained.

Chapter 2 reviews recent developments in the targeted chromatography–mass spectrometry analysis of biologically relevant endogenous carboxylic acids, addressing specific issues for small organic acids, fatty acids, eicosanoids, and bile acids. Sample preparation, derivatisation techniques, separation and MS detection of these different carboxylic acid classes are evaluated. Ultimately, based on structural features, the reader is guided to the most versatile, sensitive and facile analytical methods for the carboxylic acid class under evaluation.

Chapter 3 evaluates different column chemistries for untargeted urinary metabolic profiling in combination with different mobile phases compositions for fast LC–MS (x min per run). Three porous HILIC materials were investigated, next to core-
shell C18-, XB-C18- and PFP-RPLC material. Standards of common urinary metabolites and pooled urine samples were examined, respectively. For evaluation of all chromatographic peaks, a peak scoring algorithm was applied, taking several quality criteria, such as retention time, peak shape and peak height, into account.

Chapter 4 describes the mild and selective labelling of malondialdehyde (MDA) with 2-aminoacridone (2-AA). Normally, urinary MDA levels are assessed by a thiobarbituric acid (TBA) assay, which suffers from poor selectivity, causing false positive results in some cases. Labelling with 2-AA is performed via a mild reaction (90 min at 40 °C in an aqueous citrate buffer), giving high selectivity with a limit of detection of 1.8 nM when analysed by fluorescence detection. This also enables very fast LC separation of only 5 min per run.

Chapter 5 explicates a novel derivatisation strategy using the empirically selected N-methyl-2-phenylethanamine as derivatisation reagent with a carbodiimide as (activating) co-reagent, for the selective derivatisation of carboxylic acids, such as the di- and tri-carboxylic acids of the TCA cycle. This procedure enables analysis of the derivatives using on-line solid-phase extraction and RPLC in combination with sensitive positive-ion ESI-MS. Detection limits range from 12 to 1000 nM, depending on the analyte. Also, the potential of the methods to analyse isotopologues is shown.

Chapter 6 presents a comprehensive GC–MS bases targeted analytical platform for the simultaneous quantitative analysis of fatty acids and sterols. Also, the possibility to analyse the isotope patterns from lipids with incorporated $^{13}$C, extracted from 2-$^{13}$C-acetate incubated cells, was evaluated. The approach is based on a sequential one-pot derivatisation using MtBSTFA and BSTFA. The validated method features short run times, straightforward sample pre-treatment allowing the
Scope

analysis of both free and bound lipids and high sensitivity showing lower limits of quantification in the low ng/mL range.