Nutrimetabolomics:
An Integrative Action for Metabolomic Analyses in Human Nutritional Studies

Appendix 2

Chapter 5 - Chromatography

Marynka M. Ulaszewska[1,†], Christoph H. Weinert[2,†], Alessia Trimigno[3,†], Reto Portmann[4,†], Cristina Andres Lacueva[5], René Badertscher[4], Lorraine Brennan[6], Carl Brunius[7], Achim Bub[8], Francesco Capozzi[3], Marta Cialiè Rosso[9], Chiara E. Cordero[9], Hannelore Daniel[10], Stéphanie Durand[11], Bjoern Egert[2], Paola G. Ferrario[8], Edith J.M. Feskens[12], Pietro Franceschi[13], Mar Garcia-Aloy[5], Franck Giacomoni[11], Pieter Giesbertz[14], Raúl González-Domínguez[5], Kati Hanhineva[15], Lieselot Y. Hemeryck[16], Joachim Kopka[17], Sabine Kulling[2], Rafael Llorach[5], Claudine Manach[18], Fulvio Mattivi[1,19] Carole Migné[11], Linda H. Münger[20], Beate Ott[21,22], Gianfranco Picone[3], Grégory Pimentel[20], Estelle Pujos-Guillot[11], Samantha Riccadonna[13], Manuela J. Rist[8], Caroline Rombouts[16], Josep Rubert[1], Thomas Skurk[21,22], Pedapati S. C. Sri Harsha[6], Lieven Van Meulebroek[16], Lynn Vanhaecke[16], Rosa Vázquez-Fresno[23], David Wishart[23], and Guy Vergères[20]

[1]Department of Food Quality and Nutrition, Fondazione Edmund Mach, Research and Innovation Centre, San Michele all'Adige, Italy
[2]Department of Safety and Quality of Fruit and Vegetables, Max Rubner-Institut, Karlsruhe, Germany
[3]Department of Agricultural and Food Science, University of Bologna, Italy
[4]Method Development and Analytics Research Division, Agroscope, Federal Office for Agriculture, Berne, Switzerland
[5]Biomarkers & Nutrimetabolomics Laboratory, Department of Nutrition, Food Sciences and Gastronomy, XaRTA, INSA, Faculty of Pharmacy and Food Sciences, Campus Torribera, University of Barcelona, Barcelona, Spain. CIBER de Fragilidad y Envejecimiento Saludable (CIBERFES), Instituto de Salud Carlos III, Barcelona, Spain
[6]School of Agriculture and Food Science, Institute of Food and Health, University College Dublin, Dublin, Ireland
[7]Department of Biology and Biological Engineering, Food and Nutrition Science, Chalmers University of Technology, Gothenburg, Sweden
[8]Department of Physiology and Biochemistry of Nutrition, Max Rubner-Institut, Karlsruhe, Germany
[9]Dipartimento di Scienza e Tecnologia del Farmaco Università degli Studi di Torino, Turin, Italy
[10]Nutritional Physiology, Technische Universität München, Freising, Germany
[11]Plateforme d'Exploration du Métabolisme, MetaboHUB-Clermont, INRA, UNH, Université Clermont Auvergne, Clermont-Ferrand, France
[12]Division of Human Nutrition, Wageningen University, Wageningen, The Netherlands
[13]Computational Biology Unit, Fondazione Edmund Mach, Research and Innovation Centre, San Michele all'Adige, Italy
[14]Molecular Nutrition Unit, Technische Universität München, Freising, Germany
[15]Institute of Public Health and Clinical Nutrition, Department of Clinical Nutrition, University of Eastern Finland, Kuopio, Finland
[16]Laboratory of Chemical Analysis, Department of Veterinary Public Health and Food Safety, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium
[17]Department of Molecular Physiology, Applied Metabolome Analysis, Max-Planck-Institute of Molecular Plant Physiology, Potsdam-Golm, Germany
[18]INRA, UMR 1019, Human Nutrition Unit, Université Clermont Auvergne, Clermont-Ferrand, France
[19]Center Agriculture Food Environment, University of Trento, San Michele all'Adige, Italy
[20]Food Microbial Systems Research Division, Agroscope, Federal Office for Agriculture, Berne, Switzerland
[21]Else Kröner Fresenius Center for Nutritional Medicine, Technisal University of Munich, Munich, Germany
[22]ZIEL Institute for Food and Health, Core Facility Human Studies, Technical University of Munich, Freising, Germany
[23]Departments of Biological Sciences and Computing Science, University of Alberta, Edmonton, Canada

†First authors

Mol Nutr Food Res DOI: 10.1002/mnfr.201800384

5. Chromatography

5.1. Organization of Samples for Chromatography

The purpose of an untargeted nutritional metabolomic analysis is to find out whether significant differences between the samples of different study groups or time points exist. While the main focus is naturally on the analysis of the actual study samples, further types of (auxiliary) samples are needed in order to ensure analytical quality and to achieve optimal results. In addition to the fundamentals delineated in sections devoted to the management of samples (chapter 3, section 4.1, section 4.2), some specific remarks about the different samples, which are more closely related to the chromatographic procedure, are provided in this section.

5.1.1. Study Samples

Study samples are all samples directly resulting from the designed nutritional metabolomic experiment. These samples are usually analyzed once for at least two reasons. On one hand, especially in case of large-scale studies and higher sample numbers, it is often not feasible to analyze all samples in duplicate or triplicate. On the other hand, from a statistical point of view, technical replicates are less useful because they cannot be treated as independent samples and have thus to be combined prior to data analysis. Accordingly, the Chemical Analysis Working Group of the Metabolomics Standards Initiative also underlines that "biological replicates are preferred over analytical replicates as biological variance almost always exceeds analytical variance".[1]

Study samples should be run at random, as contamination will gradually build up in the MS source leading to changes in sensitivity. Therefore, randomization will ensure that all the experimental groups are affected by instrument response variability to the same extent. This can be achieved by using a randomized block design, whereby samples are classified into relatively homogeneous subunits containing the desired experimental design. This will reduce noise and variance of the data.[2, 3] More information about randomization is given in section 3.6.2.

5.1.2. Quality Control Samples

During untargeted analysis, the analytical variability needs to be minimized as far as possible, which is achieved by following standardized procedures for sample preparation and instrumental analysis. Nevertheless, a certain part of the analytical variability is unavoidable and needs to be tightly controlled. For this reason, Quality Control (QC) samples have commonly been used in order to monitor the metabolomic workflow. QC samples are typically composed of pooled study samples or of a representative subset of subjects in the case of larger studies.[3] QC samples contribute to the effective monitoring of the performance of the analytical system, especially in terms of absolute t_R (column conditioning), signal intensity and mass accuracy. Additionally, QC samples need to be injected at regular intervals during the batch to assess repeatability and to enable a correction of instrumental drift (see section 5.1.5). More details about preparation and meaning of QC samples are given in section 3.6.1.

5.1.3. Reagent Blanks

Reagent blanks are prepared exactly like the study and QC samples with the exception that the sample matrix is replaced by, e.g., the extraction solvent. The reagent blank is needed to enable the identification all kinds of artefacts that are present in the extraction solvent, the derivatization reagents, contaminants leaching into the samples from the labware, or reaction byproducts that are formed during the derivatization reactions. The reagent blank should be prepared for every batch to enable the detection of differences between the batches.

Moreover, extraction solvent blanks, i.e. solvent only, are also used in analytical run and serve to filter out contaminant peaks by 'blank subtraction' during data processing.[4] Additionally, they are injected at the end of the run for column washing, together with a strongly eluotropic solvent, i.e. ACN or MeOH.[3]

5.1.4. Validation Samples

In order to evaluate the effectiveness of any normalization procedure, further validation samples are required. These samples should be matrix samples resembling the study samples; for example, they can be prepared from the QC sample pool or using representative study samples. If analyzed repeatedly like the QC samples, but without being considered for the construction of the QC sample-based correction functions, data obtained with validation samples can be used to monitor the results of the different data processing steps and to estimate the effect of the normalization technique on the study samples.[5]

5.1.5. Sample (Measurement) Order

Designing an appropriate sample order is critical both for GC and LC. A specific feature of GC-MS metabolomics is the need to increase the volatility of the analytes by derivatization in order to enable GC separation (see section 4.1.3). Reports concerning the stability of the (methoxime-) trimethylsilyl derivatives[6-8] indicate that a partial degradation of derivatives may occur over time which highlights the need to analyze the derivatized samples as soon as possible, usually within 24 hours. For this reason, the samples have to be prepared freshly every day and a GC-MS metabolomic measurement is typically organized in day-wise batches.[9] Only if the derivatization can be automated,[10] a more continuous measurement process and thus longer batches are possible. In case of LC-MS metabolomics, a derivatization is not necessary and the stability of the samples in the (typically cooled) autosampler is commonly considered as less critical which enables longer analysis sequences.

Concerning the measurement order within and beyond these batches, the following recommendations may be given for both GC-MS and LC-MS techniques:

- *Equilibration runs:* An untargeted metabolomic analysis is always associated with the continual injection of matrix-burdened samples which progressively contaminate central parts of the GC or LC system, i.e. the injection port, the column and the mass spectrometer. This may result either in an increased or decreased response; while some analytes can only be detected if they are injected together with a complex matrix (i.e., matrix enhancement), the peak shape and the limit of detection of other compounds are

severely affected by a progressive contamination of the analytical instrument. Due to this fact, MS-based analytical platforms[11] need to be equilibrated with matrix samples in order to achieve stable signal intensities. This is especially necessary after installation of a new column, at the beginning of the measurement series and after a break (e.g. weekend) or maintenance (liner exchange, cleaning of ion source). Running several QC samples is probably the best way to equilibrate a GC-MS/LC-MS system as these samples are closest to the study samples. According to literature recommendations[9, 11, 12] and our own experience, between 4-6 equilibration runs with QC or other samples of the same matrix are needed to equilibrate a LC-MS or GC-MS system. Further, if the septum of a GC-MS system has to be replaced within a batch, it is advisable to continue with one or two QC equilibration runs to avoid short-term signal fluctuations.

- *Reagent blanks or standard mixtures:* These samples contain no matrix and have not the same 'equilibrating' effect as the injection of matrix samples. In contrast, the injection of a single reagent blank (see section 5.1.3) does not affect signal intensities of an already equilibrated GC/LC system. For this reason, we recommend to inject the daily reagent blank at the beginning of each batch and, if applicable, after the equilibration runs.
- *QC samples:* Generally, QC samples should initiate and terminate each block of matrix samples within a batch. Literature recommendations concerning the frequency of the QC sample injections are, however, different for GC-MS- and LC-MS-based metabolomics: For GC-MS metabolomics, it has been recommended to inject one QC sample after 4-5 study samples.[12, 13] For LC-MS metabolomics, QC sample injections are typically performed in duplicate after every batch of ten samples and the average signal of those two injections are used to normalize the ten preceding samples.[3, 14, 15] Kamleh et al.[14] observed a slightly better normalization when higher rates of QC samples were injected, but obviously, this resulted in longer analysis time that should be avoided at any cost in high-throughput analysis.[14]
- *Study samples*: Besides the arrangement of study samples in blocks flanked by QC samples, their randomization strategy is relevant. In principle, two randomization schemes are possible: complete randomization and partial or group-wise randomization. (More about randomization is given in section 3.6.2). The choice of the randomization scheme may thus depend on the research question.

5.2. Injection of GC Samples

The injection of the sample into the injection port is a decisive step of the GC analysis. It serves the purpose to introduce the (typically) liquid sample into the injection port and to transfer the analytes into the gaseous mobile phase that enters finally the column. Generally, the injection and evaporation process is influenced by a large number of factors[16, 17] and is therefore a potential source of different kinds of errors that may occur during a metabolomic analysis. Accordingly, in this section, practical recommendations concerning the automated injection using an autosampler will be provided and several aspects related to the different injection techniques will be discussed in detail.

5.2.1. Autosampler and Injection Syringes

To render the injection and evaporation process efficient and reproducible, a high degree of standardization is required which can only be achieved using automated injection. To ensure a long-term optimal performance of the autosampler, the syringe, which is the core piece of the autosampler, needs regular inspection and cleaning as it is frequently in contact with the chemically aggressive, and matrix-burdened, derivatized samples. Further, some autosampler parameters need to be chosen appropriately to enable an optimal chromatography and to achieve reproducible results. **Table S5.1** gives an overview of the most critical points.

Several injection techniques have been developed in recent decades, all of which are appropriate in certain situations. In general, also in the metabolomic field, predominantly the classical split/splitless injection is the technique of choice, with hot and cold injection being the main varieties. In view of the broad application of this technique in the metabolomic community, its characteristics are briefly discussed in this section.

In case of split injection, most of the sample introduced into the injection port is vented off through the split exit while just a minor part is passed to the column. On one hand, this means that detection limits are rather high. On the other hand, the comparatively high gas flow through the injector (at least ca. 8-10 mL/min, depending on the split ratio) ensures a fast transfer of the analytes from the injection port to the analytical column, minimizing the risk of adsorption or degradation of analytes in the injection port. In case of splitless injection, the split valve is closed before the injection and the entire sample is thus transferred to the column which ensures low detection limits. In contrast, the much lower total gas flow causes a slow transfer of the analytes to the column, which means a higher risk of adsorption or degradation in the injector. After a period of ca. 0.5-2 min (the so-called splitless hold time), the split valve is opened again and the remaining sample material is removed from the injector mainly through the split line. Consequently, split injection is typically used for more concentrated samples while splitless injection is adequate for more diluted samples In a metabolomic study, abundant as well as trace analytes are potentially of interest and the analytical method should be developed in a way that as many analytes as possible can be detected within the linear dynamic range of the detection system. In principle, this can be achieved in two ways:

- If comparatively small volumes of the derivatization reagents are added to the dried sample material, the derivatized samples are rather concentrated and split injection has to be used. This practice is very common in the metabolomic field, probably due to the fact that split injection is easy to use and needs almost no optimization. However, this approach can have a major drawback: a small excess of the derivatization reagents may results in the formation of relevant amounts of incompletely derivatized artefacts that may interfere with analytes of interest (reduced intensity of, and sometimes coelution with, the 'right' derivatives) and, in the worst case, contaminate the analytical system. Prominent examples are the non-methoximated but silylated sugar derivatives that can usually be observed in matrices with high sugar content like plasma or serum and whose formation is often difficult to avoid completely.[18, 19]

Table S5.1. Critical points on autosampler and injection syringes.

Parameter	Recommendation
Inspection and cleaning	During a measurement series, check syringe once a day for ease of motion and visual contaminations. If needed, rinse the syringe body and the needle several times with solvents like acetone and heptane. Wipe the plunger and the syringe needle with a lint-free cloth and the aforementioned solvents. Brownish deposits (by-products of the derivatization reagents) can be removed by rinsing with 50 mM hydrochloric acid (or 50 mM sodium hydroxide), water and finally acetone. Finally, rinse the syringe several times with (fresh) washing solvents after re-installation in the autosampler. According to the author’s experience, syringe cleaning is more frequently required if plasma/serum samples are analyzed compared to other matrices like urine
Washing solvents	At least two different solvents should be used to rinse the syringe before and after the injection. The use of protic solvents like methanol is not recommended as they react readily with silylation reagents, accelerating the formation of deposits in the syringe body. Instead, we recommend the use of acetone and heptane (0/3 and 6/6 washing cycles before/after injection, respectively). During each washing cycle, the rinsing volume should be at least 50% of the total syringe volume, better more. Washing solvents should be replaced and the reservoirs cleaned or replaced at least after 3-4 days or as soon as particles and precipitates can be observed. In case of matrices with a higher lipid or sugar content like serum/plasma, a daily change of the solvents is necessary to avoid carry-over
Syringe volume	The use of 10 µL syringes is recommended. While a slightly higher injection precision may be achieved using 5 µL syringes, their plunger is more filigree and thus more prone to bending when deposits accumulate in the syringe body
Needle type:	Syringes with fixed or with removable needles are available. Both are appropriate but the latter have the advantage that they can be repaired if the needle has been bended during an autosampler error
Needle tip	A cone-shaped needle tip is advantageous as it improves septum lifetime compared to bevel tip needles
Plunger type	Syringes with PTFE-tipped plungers are preferred as they have a higher durability. Further, PTFE-tipped plungers can be replaced
Filling speed	Draw up the sample liquid into the syringe body not too quickly due the viscosity of the sample liquid. 2-5 µL/s is a reasonable speed; a pull-up delay of up to 500 ms may be added
Filling strokes	Usually, 2-3 strokes are required to remove air bubbles from the syringe body after drawing up the sample liquid
Injection speed	A slow injection may cause bad peak shapes of early-eluting compounds. For example, 50-100 µL/s are appropriate
Post-Injection delay	Keeping the syringe needle in the injection port after injection too long may cause irreproducible results. A post injection delay of 200-300 ms is usually a good choice
Injection modes	Numerous injection modes (also referred to as 'needle handling techniques'[16] in the era of manual injection) have been applied in the past. According to the author’s experience, a simple 'sample only' injection works well for metabolomics samples. For example, drawing up a small air volume (e.g., 0.5 µL) after the sample liquid has no measurable effect on injection precision.

- Another approach is to add the derivatization reagents in a higher excess. In this case, the formation of incompletely derivatized artefacts is suppressed, but the final derivatized samples are much more diluted which necessitates splitless injection. The main problem with this injection technique is that it requires optimization in order to achieve a sufficient analyte response, adequate peak shapes and a good separation, especially concerning the smaller and more volatile analytes. The splitless hold time and the initial temperature of the GC oven are among the most critical parameters that need to be optimized. Apart from this, other factors like the boiling point of the sample solvents/derivatisation reagents, the injection port liner type, the injection port temperature, column dimention, or the gas controlmode during the splitless hold time (i.e., normal vs. high-pressure injection) need to be considered as they also may have a major impact on the result. Further, if metabolomics samples are injected in splitless mode, an initial GC oven temperature of about 50-70°C may be preferred to enable a proper focusing of the derivatives in the column head and thus an adequate peak shape.[20-22] Moreover, as the complete sample is directed to the column, splitless injection may also result in a faster contamination of the capillary column by matrix components.

In summary, in contrast to the widespread use of the split injection technique in the metabolomic field,[9, 12, 23-28] the application of splitless injection has the potential to overcome some major problems associated with split injection but comes at the cost of higher optimization efforts. A detailed comparison of both techniques for the analysis of metabolomic samples is, to our knowledge, still missing (see also[29]).

Another relevant question is whether hot or cold injection should be preferred. Again, both approaches have their pros and cons:

- Hot injection is undoubtedly the standard technique in the metabolomics field.[9, 12, 20, 21, 24-27, 30] The term 'hot' refers to the high injection port temperature of typically 220-280°C which is maintained throughout the analytical run. Upon injection, the sample liquid evaporates rapidly and forms a vapor/aerosol cloud. The volatile analytes are passed on to the column more or less quickly, depending on their volatility. The non-volatile components deposit in the injector liner or are in part transferred to the column head via aerosol droplets formed during the explosive evaporation of the solvents/reagents. Hot injection is easy to use and requires no specific hardware as it can be performed with every standard split/splitless injector. Major drawbacks of the hot injection technique are thermal decomposition of labile analytes due to the sudden heating and the low response of high-boiling compounds due to their slow and thus incomplete vaporization in the injection port (also referred to as discrimination[31]).
- Cold injection, also referred to as programmed temperature vaporization (PTV), aims to overcome the drawbacks of hot injection by introducing the sample liquid at low temperatures (usually below 100°C) and vaporizing them softly by rapid heating afterwards. It is commonly assumed that the slower heat transfer reduces the risk of thermal decomposition of labile analytes. The discrimination of high-boiling analytes and, at the same time, the unwanted transfer of non-volatile contamination via aerosols droplets is reduced. Moreover, PTV injectors offer a considerable flexibility concerning the development of specific temperature/split ratio programs that enable, e.g. the removal of high-boiling compounds through the split line after passing on the volatile compounds to the column or also the injection of large sample volumes.[17, 31] However, a specific hardware and an optimization of additional parameters are necessary. Due to the longer residence time and therefore the potentially increased contact of the analytes with the glass or metal surfaces in the injection port, active compounds may be adsorbed irreversibly.

Although a detailed comparison of hot and cold injection techniques for metabolomic analyses is still missing, it can be observed that cold injection becomes more and more popular which may be an indicator of positive experiences made by the analysts in everyday labwork.[22, 28, 32] Using PTV injection, Weinert et al.[28] performed more than 500 injections of urine samples without observing any signs of peak shape deterioration and thus without the need to trim the (pre-) column.

5.2.2. Injection Port Liner

An indispensable functional part of the injector is the injection port liner, a hollow glass tube in whose inward the evaporation of the sample liquid takes place. Basically, this part has the function to transfer the heat from the injector metal block to the sample liquid and to retain non-volatile sample components in order to protect the capillary column against contaminations. The type of liner has to be chosen according to the injection technique: Split liners usually have a large internal diameter (ID) and a large inner surface to enable a fast and complete mixing of the sample vapor cloud with the carrier gas. Splitless liners have a small ID in order to speed up the transfer of the analytes to the column at low gas flow rates. The use of glass/quartz wool plug at a fixed position in the liner may be advantageous for two reasons: On one hand, a wool plug increases the evaporation surface in the liner further and helps to speed up the formation of a homogeneous vapor cloud. On the other hand, wool helps to retain particles and non-volatile contaminants in the liner.
While the type of liner used is often not reported in metabolomics papers at all, the use of liners filled with wool has been recommended repeatedly.[12, 22, 28, 30] Both the liner and the wool should be chemically deactivated to remove active silanol groups at the glass surface which may reduce analyte recovery or cause degradation or a bad peak shape. To ensure a complete and reproducible deactivation of the liner and the wool, it is generally advisable to use commercially deactivated, pre-filled liners. It is unclear if the different liner deactivation methods offered by commercial suppliers have a major impact on analytical parameters like recovery/response, peak shape, and precision in case of an untargeted metabolomic analysis because investigations in this respect are missing.

The permanent injection of highly matrix-burdened samples means anyway a progressive contamination of the liner and thus the formation of new active sites, independent of those potentially on the glass/wool surface. Further, it can be assumed that at least the active sites in the injector– pre-existing or newly-formed – that contain acidic protons are neutralized by the permanent injection of silylation reagent.

Combining all this, it seems to be more important to keep in mind that the liner is a consumable that needs to be exchanged regularly but the exchange intervals depend strongly on the sample matrix. In case of urine and plasma/serum, the liner should be exchanged at the beginning of the week or after 120-150 matrix sample runs, at least if split injection is used (**Box S5.1**)

5.2.3. Injection Port Septum

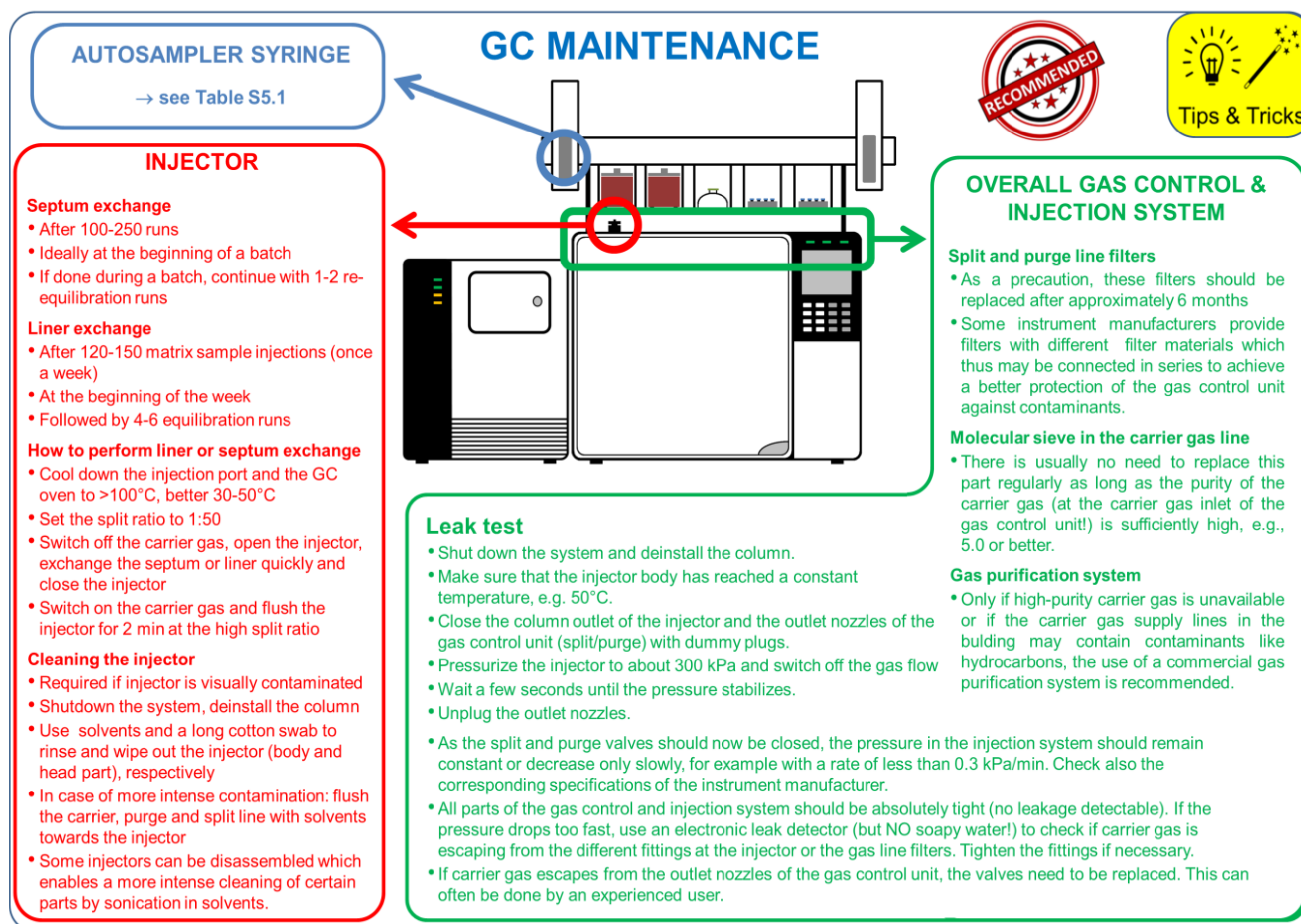
Like the liner, the injection port septum is an important consumable that requires some attention. The septum is penetrated during each injection which causes – sooner or later – leakages and an influx of air. Further, detached septum particles may fall into the liner. A regular exchange is therefore needed to avoid these problems but the exchange intervals may vary according to the septum material, the tip type of the syringe needle (see section. 5.2.1 and **Table S5.1**) as well as the injection technique. We recommend pre-conditioned Thermolite® septa from Restek® due to their durability. These septa usually withstand up to 250 or approx. 100 injections in case of hot or cold injection, respectively. It has to be kept in mind that changing the septum means to open the injector and to allow a certain influx of air into the system; this triggers multiple chemical reactions between oxygen and all kinds of compounds in the injection port and the column (stationary phase, surface deactivation, contaminants, etc.), which change the overall physicochemical condition– and may cause the GC-MS system to depart from the 'equilibrated' state. An appropriate procedure for septum and liner exchange as well as further tips for GC maintenance can be found in **Box S5.1**.

5.2.4. Injection Volume and Expansion Volumes

Generally, injection volumes in GC-MS analysis are usually restricted to about 1 µL due to the fact that the sample liquid expands dramatically upon evaporation in the hot injector. Therefore, injecting too large volumes of the sample liquid

means to produce a vapor cloud that exceeds the internal volume of the liner (split liners have typically an internal volume of ca. 500-800 µL). As a consequence, sample components would be expelled out of the liner and contaminate the septum region of the injector as well as the carrier gas and the septum purge line, which may have all kinds of spurious effects like ghost peaks, carry-over, etc. For this reason, it is vital to compare the vapor volume of the injected volume of the respective sample liquid under injection conditions (pressure and temperature in the injection port) with the internal volume of the respective liner. A rule of thumb is that the vapor cloud should not exceed 75% of the internal volume of the liner. In case of GC-based metabolome analyses, the sample 'solvent' is usually a mixture of pyridine and silylation reagents like MSTFA or BTSFA; especially the latter have a higher molecular weight than the classical solvents (199.3 and 257.4 g/mol, respectively). As can be deduced from the ideal gas law, the (expansion) volume of a solvent is inversely proportional to the molecular weight. In principle, the solvent's density is another major factor; however, with the exception of the chlorinated solvents and the alkanes, most solvents have a density between 0.7-1.0 g/cm^3, which means that density has in practice a limited influence on expansion volumes.

Overall, this means that the expansion volume is comparatively small if the sample liquid consists predominantly of MSTFA or BSTFA, which is usually the case. As a consequence, injection volumes of up to 2.5 µL are theoretically possible (see **Table S5.2**), assuming the use of an appropriate (split) liner with a sufficient internal volume. Further, it is worth to note that the expansion volume can be lowered by decreasing inlet temperature or increasing inlet pressure. This means that the use of PTV injection and high pressure injection also enable higher injection volumes. However, as higher injection volumes may lead to broader peaks, a reasonable compromise between increased sensitivity and acceptable peak widths needs to be found. While an injection volume of 1 µL seems to be standard in the metabolomics literature,[9, 12, 20-22, 24, 25, 28, 32] also injection volumes of 2-3 µL have been reported.[23, 26, 27]



Box S5.1. GC-MS maintenance.

Table S5.2. Expansion volumes of different volumes of a mixture of 15 µL MeOX in pyridine (20 mg/mL) and 50 µL MSTFA at 250°C and inlet pressures of 100 kPa and 200 kPa, as determined with the 'Vapor Volume Calculator' provided by Agilent Technologies.

Injection volume (µL)	Expansion volume (µL)	
	p = 100 kPa	p = 200 kPa
0.5	74	49
1.0	150	100
1.5	226	150
2.0	302	201
2.5	378	252

5.2.5. Remarks Concerning Maintenance of the Gas Control and Injection System

Purge Flow

The carrier gas stream is divided in the injector into i) the part that is directed to the column (column flow), ii) the split flow, and iii) the purge flow, which helps to eliminate volatile contaminants from the region below the injector septum and around the head of the liner.

Derivatized metabolomic samples contain always large amounts of highly volatile compounds that are no metabolites of interest, like the derivatization reagent itself, pyridine as well as all kinds of derivatization side-products. These compounds can reach the head of the liner easily (even if appropriate sample volumes have been injected) and may lead to strong tailing of the solvent front if they are not removed appropriately. To achieve this, the purge flow must always be activated and sufficiently high, for example 6 mL/min.

Split and Purge Line Filters

In a typical GC instrument, the split and the purge line contain filters that shall protect the gas control unit (especially the valves) against organic contaminations and to prevent potentially hazardous compounds from entering the laboratory atmosphere. While the latter is especially important if toxic compounds like pesticides or polycyclic aromatic hydrocarbons are analyzed, the former is relevant if derivatized metabolomic samples are analyzed on a routine basis. The filters are usually filled with charcoal, sometimes combined with silica. It is commonly assumed that most of the organic compounds that enter the split and the purge line are trapped by these filters. However, metabolomic samples contain a large number of highly volatile compounds (the derivatization reagents, their by-products as well as the metabolite derivatives) that can form sticky deposits, for example on the inner surface of the gas lines. Experience shows that, at least after some time, these compounds can also penetrate the filters and cause malfunctions of the valves in the gas control unit. In an advanced stage, this may lead to visible deposits in the outlet nozzles of the gas control unit. In case of splitless injection, the vast majority of the volatile compounds in the sample are directed to the column and a negligible part of them is exiting the injector via the purge line, which means that the purge line is usually not greatly contaminated. In contrast, in case of split injection, most of the sample is removed from the injector via the split line, which can lead to strong contamination even after a comparatively short time. As a consequence, the function of the valves is impaired: Split and purge flow cannot be regulated properly, which affects t_R stability as well as precision. In the worst case, the valves work sometimes and sometimes not, leading to spurious random changes of overall system performance.

Leak Test of The Gas Control and Injection System

The quickest and most sensitive way to check the tightness of the entire injection and gas control system as well as the integrity of the valves in the gas control unit is to perform a comprehensive leak test (see **Box S5.1**). A possible outcome of this test may be that the valves of the gas control system have been contaminated and need to be replaced. In extreme cases, if also the gas lines within the gas control unit are strongly contaminated, the entire gas control unit must be replaced. The only way to avoid (or at least reduce) such problems is to exchange the filters in the split and purge line frequently, for example after 4-6 month or after the injection of approx. 2,000-5,000 samples, depending on several factors like the sample matrix, the injection volume, etc. At the same occasion, it is advisable to backflush the gas lines between the injector and the filters with different solvents. However, even if such a rigorous preventive maintenance scheme is applied, a long-term contamination of the gas control unit can probably not be avoided completely. For this reason, the increased application of splitless injection should be considered as another way to reduce the contamination of the gas control system by volatile organic compounds originating from the derivatized samples (see section 5.2.2).

5.3. One-Dimensional Gas Chromatography

5.3.1. Carrier Gas

In a GC system, the carrier gas acts as the mobile phase. Despite an increasing shortage of natural supplies, helium is still the carrier gas of choice because it combines several favorable properties: i) it is an inert gas that does not react with, e.g., sensitive analytes; ii) it enables comparatively fast separations; ii) it can be pumped away efficiently by the turbomolecular pumps of MS systems, enabling low baseline noise and thus low limits of detection. The next alternative is hydrogen which has the advantage that i) it can be supplied on-site in high purity using hydrogen generators and ii) it enables even faster separations than helium. In contrast, several analytes can be hydrogenated upon reaction with elemental hydrogen, especially in the hot sections of a GC-MS system.[31] As hydrogen can form explosive mixtures with air, its use as carrier gas is always connected with a safety risk. Further, turbomolecular pumps cannot remove hydrogen as effectively as helium, leading to higher baseline noise and a sensitivity drop by a factor of 2-5 compared to helium. Especially the latter fact is usually seen as a decisive drawback, especially because metabolomic experiments are usually done in full scan mode where no possibility exists to reduce baseline noise. Last but not least, nitrogen can also be used as a GC carrier gas. Although it is an inert gas like helium, its gas chromatographic properties are not optimal (in brief, good but slow separation). Therefore, nitrogen, like hydrogen, is not being considered as a realistic alternative to helium in metabolomic applications so far, despite the expected shortage of the natural helium resources.

5.3.2. Capillary Columns

The capillary column is the part of a GC system where the chromatographic separation takes place. In a metabolomic situation, this separation task is especially challenging due to the following circumstances:

- Biological samples contain high numbers of compounds that are structurally either similar or dissimilar to each other. The concentrations of these compounds may differ by several orders of magnitude.
- In addition to these potentially interesting 'metabolites', different kinds of matrix components and, in case of GC-based metabolomics, derivatization reagents as well as their side products, are present; they increase the chemical complexity of the sample mixture and may in many ways influence the separation, detection and quantification of the metabolites.
- Derivatized samples contain chemically aggressive reagents that may react, amongst others, with the stationary phase unless the latter is sufficiently inert
- Some metabolites or matrix compounds may be relatively non-volatile which necessitates a bake-out step at the end of the GC temperature program. Thus, the stationary phase needs to be temperature-stable.
- Concerning compound identification, the availability of retention indices for the respective stationary phase is critical (see sections 3.6.3 and 9.3.2)

These requirements can be directly translated into recommendations concerning the choice of the column dimensions and the stationary phase:

- Especially non-polar or medium-polar phases fulfil the criterion to be chemically inert and sufficiently temperature-stable (maximum programmed temperature 320-350°C) at the same time. For this reason, especially 100% dimethyl polysiloxane (e.g., DB1[30]) and 95% dimethyl/5% (di)phenyl polysiloxane (e.g., DB[20-22, 24, 26, 27]) are the most commonly used stationary phases in metabolomic laboratories. As a consequence, the analytes (non-polar metabolite derivatives) are predominantly separated according to their volatility. An alternative may be the use of stable mid-polar phases like BPX50 (phenyl content of 50%[9, 12]).
- The fact that the retention indices provided in most spectral libraries were primarily determined on non-polar phases also speaks for their use.
- Besides the chemistry of the stationary phase, the column length is a major determinant of chromatographic separation. However, if column length is doubled, chromatographic resolution increases only by a factor of 1.4 while analysis time is nearly doubled.[31] At the same time, it is virtually impossible to achieve a sufficient separation of all components of a very complex biological sample only by means of one-dimensional GC, even with a 60 m column. For this reason, a column length of 30 m is commonly considered as standard[9, 12, 24, 26, 27, 30, 32, 33] because it is a reasonable compromise between chromatographic efficiency and run time, bearing in mind that coeluting compounds may finally be “separated” by the MS.
- In general, there are many points that advocate the use of thin film columns: They yield narrow peaks, enable a more efficient separation, and are featured by low elution temperatures and low bleeding. Likewise, lowering the ID of the column tubing is in principle beneficial as this enables a higher chromatographic resolution (higher efficiency, i.e. more plates per meter) and a lower column flow which means a lower pumping “burden” for the turbomolecular pumps of the MS and in turn lower detection limit. However, the concomitant disadvantage of a thinner film or a lower ID is a reduced sample capacity.[31] As metabolomic methods aim to resolve trace level analytes from highly abundant compounds, an ID of 0.25 mm and a film thickness of 0.25 µm have shown to be a good compromise.[9, 12, 24, 26, 27, 30, 32, 33]
- The injection of rather crude sample extracts usually means that non-volatile components accumulate not only in the injection port liner but also in the head of the capillary column. This may affect, sooner or later, analyte peak shapes and lead to a loss of sensitivity. Deactivated pre-columns, also referred to as guard columns or retention gaps, are used to trap such contaminants in an “empty” zone of the tubing void of stationary phase which can be removed by trimming. In this respect, the use of integrated pre-columns which eliminate the need to couple a separate pre-column to the main column via a special connector is advisable as this avoids an increased dead volume and the inherent risk of leakages linked to the use of column connectors.

5.3.3. Temperature Program

In contrast to traditional targeted analyses, not much attention is usually paid to the optimization of the GC temperature program in case of untargeted metabolome analyses. This has much to do with the fact that a complete chromatographic separation of hundreds of analytes is obviously impossible and the resulting common perception of the chromatography as just a “preparatory step” that paves the way for the final MS detection. Consequently, it is difficult to provide clear recommendations concerning the choice of the GC temperature program as this depends highly on the value one attaches to this step. However, the following general recommendations may be appropriate:

- The initial temperature of the GC temperature program should usually be about 50-70°C[9, 12, 20, 21, 24, 26, 27, 30, 32, 33] but no more than 90°C[28, 34] due to the need of re-condensing the solvent at the head of the capillary column and take all benefits from the band focusing effect. In addition, using a higher initial temperature would make it difficult to separate early-eluting analytes like the TMS derivative of pyruvate from the solvent/reagent front.
- The slope of the temperature ramp is typically chosen according to the emphasis the analysts puts on the chromatographic separation. If a high sample throughput is desired, the temperature of the GC oven may be increased at a rate of 10-30 C/min;[9, 12, 20, 21, 26, 27, 30, 32, 35] If an overall better separation of the analytes shall be achieved, a shallower ramp (5-10 C) is recommended.[22, 24, 33] While additional temperature plateaus may help to improve the separation of certain critical pairs of analytes further, every change in the slope of the ramp may affects precision of the determination of the retention indices of the analytes. Thus, the use of a rather simple temperature program is generally advisable.
- The final temperature lies usually between 290-330°C,[9, 12, 20, 21, 26, 27, 32, 35] depending on the occurrence of rather non-volatile analytes like saccharides or sterols. In most cases, it is necessary to hold the final temperature for some

minutes (bake-out step) in order to avoid carry-over between subsequent runs.

- According to recommendations from column manufacturers, the interface or transfer line temperature should preferably be kept below 300°C. If this final segment of the capillary column is permanently kept at higher temperatures, the following may happen: Firstly, the stationary phase may be gradually removed here due to the combination of high temperatures and the permanent MS vacuum, leading to the disclosure of active sites on the surface of the tubing. Secondly, after some time, higher temperatures may "burn" the polyimide coating (visual browning!) and cause the underlying tubing to become brittle, which may finally lead to a rupture of the column in the interface. For these reasons, an interface temperature of 280-300°C is recommendable, especially in the long term.

5.4. Comprehensive Two-Dimensional Gas Chromatography (GC×GC)

5.4.1. Overview

An important methodological aspect in metabolomics is to enable a selective detection and quantification of as many analytes as possible. In case of the one-dimensional GC-MS approach, due to the often incomplete chromatographic separation, the MS plays a major role in achieving the desired level of selectivity as it enables a distinction between coeluting analytes. In this respect, the use of high resolution instruments is nowadays almost considered as mandatory. In contrast, an excellent selectivity can already be achieved at the chromatography stage by means of comprehensive two-dimensional gas chromatography (GC×GC). The fundamental principle of this advanced chromatographic technique is the separation of all analytes in a given sample using two different separation mechanisms, i.e. two columns with different stationary phases. GC×GC combined with MS detection (GC×GC-MS) has been introduced in the metabolomic field some 15 years ago; since this time, GC×GC-MS has developed into an increasingly popular alternative to the common one-dimensional approach. The theory and the technical aspects of GC×GC(-MS) have been outlined exhaustively elsewhere[36-42] and shall therefore not be covered here in detail. Instead, this section aims to provide some practical recommendations concerning the use of GC×GC-MS for untargeted metabolomics, in addition to the aspects covered in section 5.3. Nevertheless, for a better understanding, the functionality of a GC×GC-MS system shall be explained briefly.

Basically, there are two things that distinguish a GC×GC-MS system from a conventional GC-MS system: The use of two serially connected capillary columns and the presence of the so-called modulator located at the interface of the two columns. In most cases, the column combination consists of a long, usually non-polar first column with common dimensions and a short, typically polar or medium-polar fast GC column in the second dimension. After injection, many analytes are already separated on the first column ("first dimension") according to the first separation mechanism (e.g. volatility) while others coelute and reach the head of the second column simultaneously. The modulator traps continually fractions eluting from the first column, refocuses and finally releases them onto the second column where the next separation step takes place.[43] The operation cycle of the modulator takes typically between 2-8 seconds[44] and the term "modulation period" has been coined for this interval. The action of the modulator has three major consequences:

- As the width of the peaks eluting from the first column is usually larger than the modulation period, all first-dimensional peaks are divided into several fractions. This means that finally multiple (usually three to five) second-dimensional peaks per analyte are generated, which have to be matched and merged during data processing.
- As the next fraction is loaded onto the second column after just 2-8 seconds, the separation on the second column has to be very fast in order to avoid interferences between subsequent fractions ("wrap-around"). For this reason, very short columns with small IDs and thin films are used in the second dimension.
- In case of a cryogenic modulator, the bands eluting from the first column are also refocused, i.e. concentrated before being reinjected onto the second column. This band focusing in-space leads to very narrow second-dimensional peaks and an improvement of the signal/noise-ratio which means in turn lower limits of detection.

As the final result of the two-dimensional separation, a much higher part of the analytes reaches the MS as pure fractions as compared to a one-dimensional setup. In the end, this means that a higher number of compounds can be reliably identified and quantified.[45] However, this comes at the cost of higher demands towards the performance of the MS (see section 6.1.2) as well as a more complex raw data structure and usually a higher data volume per run (see section 7.1.6).

The major challenge with GC×GC separations of metabolomic samples is that it is not enough to obtain structured chromatograms with roughly separated groups of compounds belonging to the same class (as in case of the well-known GC×GC separations of petrochemicals[46]) or to separate a few relevant analytes from a hump of less interesting matrix components. In case of metabolomics, essentially every single sample component is of interest. Consequently, untargeted GC×GC-MS metabolomic approaches aim to achieve a nearly-optimal separation in both dimensions in order to enable the separation and (relative) quantification of as many (chemically highly different) metabolites as possible. In addition, further aspects, among others, play a major role in routine GC×GC-MS metabolomic analysis:

- Extreme concentration differences between analytes or analyte classes.
- Long-term precision.
- Effects of progressive contamination of the system due to permanent injection of matrix-burdened samples.
- Robustness, ease of use and running costs of the instrumental platform.

Although several excellent, in part exhaustive, review articles about the GC×GC methodology have been published in recent years,[38-45, 47-49] the specific requirements and practical challenges of GC×GC-MS metabolomics have, nevertheless, not been addressed sufficiently so far. For this reason, the scope of this section is to highlight the specific challenges of GC×GC-MS metabolomic applications and, based on this, to provide practical recommendations to optimize the methodology. In addition, **Box S5.2** shows

typical GC×GC-MS chromatograms of derivatized urine and serum/plasma samples and explains the practical results of the two-dimensional separation in the light of the GC×GC essentials briefly explained in this and the following subsections.

5.4.2. Column Setup

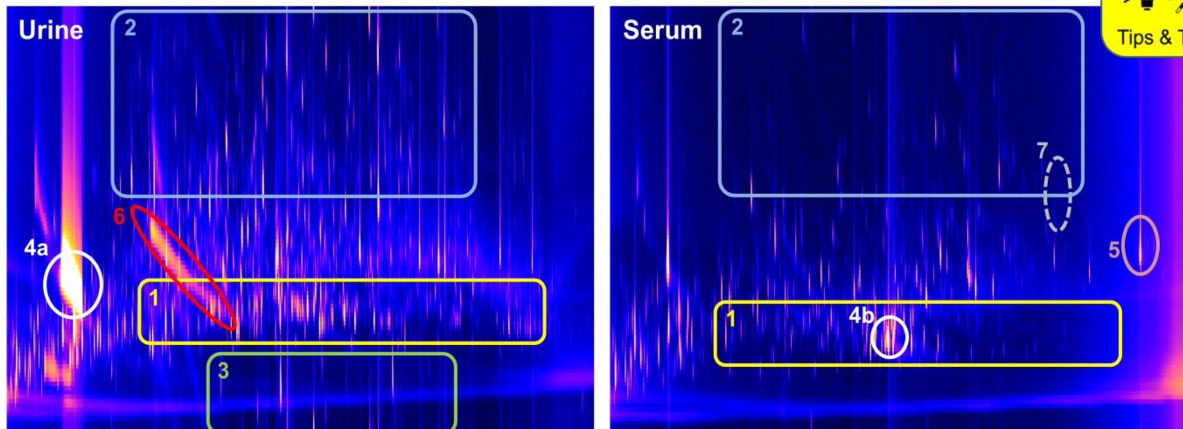
A general principle of two-dimensional separations is that the separation mechanisms in the two dimensions should be as independent as possible in order to exploit the two-dimensional separation "plane" as best as possible (principle of orthogonality). In view of the currently available stationary phases, this means that basically two types of column combinations are possible: non-polar × (medium-) polar and (medium-) polar × non-polar.[39] As numerous stationary phases with different column chemistries and spanning a wide range of polarities are available, theoretically a very large number of column combinations is conceivable. However, in view of metabolome analyses, the specific criteria described in section 5.3.2 apply which limits the number of practically usable stationary phases. Especially the fact that most strongly polar stationary phases are not suitable for metabolome analyses (due to their limited thermal stability and chemical inertness) makes it nearly impossible to make use of a truly "orthogonal" column combination. In practice, the combination of a non-polar first column with a medium-polar second column has developed into a standard setup - also in the metabolomic field.[23, 25, 28, 34, 45, 50, 51] Although it is clear that the full potential of two-dimensional separations cannot be exploited with such a column setup, this combination has proven to be long-term robust[28] and to enable an improved chromatographic separation of the methoxime/ trimethylsilyl derivatives of primary metabolites (as compared to the one-dimensional approach[45]) which are probably the main reasons for its popularity. For these reasons, the classic non-polar columns which are very popular for one-dimensional GC-based metabolomics (especially DB1- and DB5-like columns) are also mainly used as first columns in GC×GC metabolomic setups.[52] Concerning column dimensions, the same compromise has to be made (see section 5.3.2) and therefore a column with a length of 30 m, an ID of 0.25 mm and a film thickness of 0.25 µm is standard. As mentioned before, medium-polar columns are preferred for the second dimension. Most popular are 50% phenyl phases (e.g. 50% dimethyl/50% (di)phenyl polysiloxane like BPX50[23, 28, 34, 51] or 14% cyanopropyl-phenyl phases like DB-1701[25, 50]).

Although the separation in the second dimension is always performed on very short fast GC columns, the exact dimensions depend to some extend on the type of modulator, on the MS detection system and other factors. The active "separation segment" of the ^{2}D column, i.e. the part where the actual 2D separation takes place, is usually 1.0-1.5 m long.[23, 25, 28, 34, 50, 51] If a cryogenic loop-type modulator is used, another 1-1.5 m is needed for the loop.[48] Concerning the ID and the film thickness, the data acquisition rate of the MS detection system has to be considered. In principle, columns with an ID of 0.1 mm and a film thickness of 0.1 µm are ideal because they enable a very fast and efficient separation. However, such columns can be easily overloaded which is especially problematic if the analytes in the respective sample matrix are present in highly different concentrations.[53, 54] Further, the resulting ^{2}D peaks are very narrow (peak base widths of ca. 50-500 ms) which necessitates the use of a fast MS detection system. Time of Flight (ToF)-MS systems enable a data acquisition rate of more than 200 Hz which makes it possible to capture the profile of such narrow peaks adequately, i.e. with at least 15-20 data points per peak. In contrast, the speed of quadrupole MS (qMS) systems is limited and their data acquisition rate depends on the scan range. As a comparatively large scan range is needed in a metabolomic application, even fast qMS systems can typically reach a maximal data acquisition rate of ca. 50 Hz (see section 6.1.2). For this reason, if only a qMS system is available, a slightly higher ID and film thickness has been recommended (e.g. 0.15 mm and 0.15 µm, respectively) in the second dimension in order to ensure a sufficient number of data points per peak.[28] Additionally, such a configuration has the benefit of a higher sample capacity and a reduced flow mismatch.[53, 55]

The serial coupling of a long conventional first column with a short and narrow second column creates the problem of highly different and non-optimal linear carrier gas velocities in both dimensions because the second column acts like a restrictor. Traditionally, the column head pressure is set so that the linear velocity is nearly ideal (or slightly too low) in the first dimension and by far too high in the second dimension. This compromise is usually accepted for pragmatic reasons but one should keep in mind that such a configuration does not enable an optimal chromatographic performance. In principle, there are two ways to reduce this problem and to achieve more similar linear gas velocities in both columns. The first is to choose rather similar IDs in both dimensions, i.e. a smaller ID in the first dimension or a larger ID in the second dimension. The drawbacks would be a smaller sample capacity of the first column or a lower efficiency (i.e. a slower and probably incomplete separation) of the second column, respectively. The second solution suggested in recent years is to divide the gas stream after the first column, either so that a major part of the gas flow is vented away[56, 57] or so that equal/defined parts of the first-dimensional eluate are directed to two parallel secondary columns with separate detectors.[50, 58] Both approaches enable much more adequate gas flows in the second dimension and thus a better two-dimensional chromatography. However, while the first necessarily means a marked loss in sensitivity, the latter is only possible with a more complex instrumental setup. Further research is necessary to find optimized column setups amenable for routine use in GC×GC metabolomics. In simple instrumental setups, the secondary column is housed in the same oven as the primary one and both columns are therefore subjected to the same temperature program. Housing the secondary column in a separate oven – either another standard GC oven or a specialized micro-oven built-in into the main GC oven – enables a specific, independent temperature program for the

GC×GC Chromatography of Urine and Serum/Plasma

Tips & Tricks



Complexity of the metabolite profile

- Using GC(×GC)-MS, a higher number of metabolites can be detected in urine compared to serum. As illustrated by the higher number of «spots» in virtually all regions of the two-dimensional urine chromatogram, this in true for most compound classes. However, the difference is most striking for the sugars and sugar-related compounds (1) and the aromatic/heterocyclic compounds (2). Consequently, in order to ensure an adequate chromatographic separation, a slower temperature program and thus a slightly longer analysis time is typically required for urine (here: 74 min (urine) vs. 66 min (plasma)).

Modulation period (P_M) and wrap around

- Both chromatograms were recorded with a P_M of 3s. Due to the higher number of compounds with a strong ^{2}D retention, the problem of wrap-around is more relevant in case of urine (3). While a longer P_M would effectively eliminate wrap-around, this would also sacrifice part of the separation achieved in ^{1}D as explained in section 5.4.3.

Deviations from spot-like peak shape (highly abundant, strongly non-volatile or degrading analytes)

- Without urease treatment, the by far largest peak in urine is formed by the urea 2TMS derivative (4a). While this compound superimposes several other analytes in one-dimensional GC, it can be separated from other metabolites using GC×GC. Nevertheless, the urea peak shows a severe ^{2}D tailing and, if the moduator loop is not perfectly adjusted, also a ^{1}D tailing. Compared to this, the glucose peak in serum/plasma (4b) is only moderately overloaded; here, the formation of non-methoximated, but trimethyl-silylated glucose artefacts (which appear in the same chromatogram region if derivatization conditions are not optimal for the sugars) are more problematic. In case of very non-volatile and thus late-eluting metabolites like cholesterol (5), the ^{2}D peak shape is nearly always impaired independent of analyte concentration. In contrast, a comet-like peak shape with ^{1}D tailing and a constantly decreasing ^{2}D retention (6) can be interpreted as a sign of degradation in the injector or on the ^{1}D column, which is typically observed for some «active» urine metabolites.

Serum vs. EDTA plasma

- Compared to serum, chromatograms of EDTA plasma exhibit at least one very large peak of an EDTA derivative whose approximate position is indicated above (7).

Box S5.2. GCxGC chromatography of urine and serum/plasma samples. The chromatograms were acquired under the following conditions: ^{1}D column: Rxi-5SilMS, ^{1}L=30 m, ^1d_c=0.25 mm, ^1d_f=0.25 µm; ^{2}D column: BPX50, ^{2}L=0.7 m (loop: 1.5 m), ^1d_c=0.15 mm, ^1d_f=0.15 µm. P_M=3 s. Detection: qMS at 33Hz, *m/z* 60-550.

secondary column and thus a fine-tuning of the two-dimensional separation.[57, 59] In a recent work, the suitability of a non-orthogonal or “reversed” setup consisting of a medium-polar first column and a non-polar second column for metabolomics was investigated.[54] According to the authors, a higher sample capacity was achieved by using ^{2}D columns with a larger ID and thicker films, resulting in a more accurate quantification as compared to an orthogonal column setup. Despite these advantages, this approach has apparently not gained a wider acceptance in the metabolomic community. This may be in part due to the following drawbacks: On the one hand, the availability of retention indices for medium-polar columns is limited which hampers compound identification. On the other hand, a medium-polar × non-polar does not enable a better separation of highly similar sugar isomers than a non-polar × medium-polar column setup and is therefore no solution to this eminent problem in GC×GC metabolomics.

As a last technical remark, especially in view of the routine use of GC×GC-MS, practical experience shows that bolted column connectors like the SilTite™ Mini Union and SilTite™ µ-Union (SGE/Trajan Scientific) or Restek MXT™ are much tighter and more reliable than conventional press-fit connectors. Once installed properly, these connectors ensure a permanent leak-free connection.

5.4.3. Modulation

As briefly described in section 5.4.1, the modulator is a critical component of a GC×GC system. The continuous process realized by this device – the collection of fractions of the first-dimensional eluate as well as their refocusing and their re-injection onto the second column – is called modulation. Different types of modulators exist and their historic development as well as their characteristics have been delineated in detail elsewhere.[39, 44, 47, 48] As the jet-based cryogenic modulators and the flow modulators are nowadays by far the most common types, their main features shall be summarized briefly.

- In case of jet-based cryogenic modulators**,** modulation is performed with the help of hot and cold gas jets. According to the literature, two subtypes are nowadays predominantly used: The dual-stage quad jet cryogenic modulator[60, 61] (commercialized by LECO) and the dual-stage loop-type modulator with two jets[62] (commercialized by ZOEX (Corporation)). In principle, nitrogen, carbon dioxide, or (dried) air are available as modulator gases, with nitrogen being considered as the modulator gas of choice and air being the low-cost option. Cryogenic modulation is still considered as the most popular approach because it enables an optimal peak shape and an effective band focusing which enables very narrow peak widths (and thus optimal sensitivity). The drawbacks of cryogenic modulation are the comparatively

high complexity of the hardware and the higher acquisition and operational costs as compared to other types of modulators.[48]

- Flow modulators are essentially valve-based modulators and a remarkable number of such modulators has been developed during recent years.[47, 48] Typically, the first-dimensional eluate is collected in a loop or chamber whose content is, after valve switching, rapidly flushed onto the second column with the help of an auxiliary gas control unit. Their advantages are i) their technical simplicity and the omission of operational costs (no modulator gases required), ii) their speed of operation which enables very fast second-dimensional separations and small modulation periods, and iii) the fact that all analytes are modulated, even the most volatile ones (no breakthrough, insufficient trapping of the analytes in the modulator…).[39, 40] The main drawback of flow modulation is the generation of very high gas flow in the second dimension (up to 21 mL/min) which usually exceeds MS pumping capacity by far. Appreciable efforts have been made to reduce this problem by fine-tuning of modulator design and conditions[63, 64] but secondary column gas flows remain to be comparatively high (at best 4 mL/min) which still compromises MS sensitivity. Further, such optimized flow modulators are still commercially unavailable. The only commercial flow modulator (provided by Agilent) still generates the usual high second-dimensional flows, making flame ionization detector (FID) detection or a more sophisticated setup with simultaneous FID and MS detection[46] the only feasible detection options.

In conclusion, cryogenic modulation is still the method of choice for untargeted GC×GC-MS metabolomics, especially because instrumental robustness, sensitivity, ease of use, and support by commercial suppliers are highly valued in routine metabolomic work. Consequently, to the best of the authors' knowledge, no reports describing the use of a flow-modulated GC×GC-MS system for untargeted metabolomics has been published so far. In the following, some practical recommendations will be provided concerning the use of cryogenic modulation in GC×GC-MS-based metabolomics.

Among the typical modulator gases mentioned above, nitrogen is usually preferred because it enables a very intense cooling und thus an effective trapping of even very volatile compounds like propane.[48] However, nitrogen as modulator gas is costly, as a consumption of 30-40 L of liquid nitrogen per day or more is not unusual. Therefore, devices like the Optimode™ modulation control unit (offered by SRA instruments), which enable, among others, a precise and time-dependent control of the cold flow can help to reduce the liquid nitrogen consumption, and thus the costs.[50] Carbon dioxide was a common modulator gas before the introduction of liquid nitrogen, but was largely replaced later as CO_2 as it is unable to trap highly volatile compounds as efficiently as nitrogen. Compressed air can be considered as the most cost-efficient modulator gas. It can be supplied everywhere with the help of a compressor and, if sufficiently dried, be cooled down to about -90°C. This temperature enables the modulation of compounds approx. as small as heptane. Air-based and thus called 'consumable-free' modulators have been brought onto the market by LECO and ZOEX, in the latter case as the air-based loop modulator ZX2.[48] In case of metabolomics, MTS derivatization transforms even small metabolites into derivatives with a comparatively high molecular weight. Consequently, even early-eluting compounds like the MTS derivative of pyruvate can readily be modulated using cold compressed air. This means that modulation with compressed air is fully satisfying for metabolomics. Only if underivatized volatile compounds instead of derivatized primary metabolites are the target compounds ('volatilomics'), the use of liquid nitrogen as modulator gas is mandatory.

As mentioned earlier, the modulation period (P_M) is the interval between the reinjection of subsequent fractions of the first-dimensional eluate onto the second column. Commonly, a range of 2-8 s is considered as adequate.[44] As a rule of thumb, the modulation period should be as short as possible and as long as necessary, due to the following circumstances:

- One of the fundamental principles of multidimensional chromatography is that compounds resolved in the first dimension must remain separated in the second dimension.[36] It is commonly assumed that this can be achieved if every first-dimensional peak is modulated at least 3-4 times (so-called 'modulation criterion').[41, 42] Practically, this means that a sufficiently short modulation period is needed in order to keep the resolution achieved in the first dimension, especially in view of closely eluting compounds with identical or very similar mass spectra that cannot be separated solely based on specific masses. A prominent example are the highly isomeric sugars which occur in human matrices like urine as well as in all kinds of plant-derived matrices and are often of interest.
- In contrast, the modulation period has to be long enough in order to enable even the compounds most retained in the second dimension to reach the detector until the next fraction is released by the modulator. Otherwise, wrap-around can occur.[42, 43, 65] In some instances, wrap-around can be tolerated unless it leads to coelution of analytes in the second dimension that were initially separated in the first dimension. The answer to the question to which extend wrap-around can be accepted also depends on the way the raw data are processed.[65]

In conclusion, these two opposite aims – a frequent sampling of the first-dimensional eluate by the modulator and the prevention of excessive wrap-around – have to be counterbalanced. Further, the choice of the modulation period depends on several factors, with i) the column combination, i.e. the stationary phases and the dimensions of both columns, ii) the slope of the GC temperature program, and iii) the composition of the sample material being among the most important. As the column combinations and the instrumental settings in recent GC×GC-MS metabolomic reports focusing on the analysis of human/animal biofluids, tissues or cells vary considerably,[13, 23, 25, 28, 34, 50, 66-68] reported modulation periods were between 2-6 s.

Finally, some practical recommendations may be given concerning the use of loop-type modulators:

- The temperature of the hot jet should increase over time, either in a linear fashion or stepwise. High hot jet temperatures at the beginning of the GC temperature program can disturb the temperature control of the GC oven and lead to inconsistent t_R. In contrast, comparatively low hot jet temperatures at the end of the GC temperature program could render band remobilization inefficient. A positive temperature difference between the hot jet and the GC oven of at least 50°C, in case of air-based modulation (e.g. ZOEX ZX2), or at least 100°C, in case of nitrogen modulation, is advisable.

- Concerning the hot jet pulse time, a compromise has to be found: Too long pulses might lead to breakthrough while too short pulses might cause incomplete remobilization. A pulse time of about 300 ms as reported by O'Hagan[23] may be appropriate in most cases.
- As the temperature of the gas exiting the cold jet nozzle (-90°C in case of air-based modulation (Corporation) or up to -190°C in case of liquid nitrogen modulation (Corporation)) cannot be controlled directly, the temperature of the cold spot can only be adjusted by controlling the cold jet flow. By default, this can be achieved via a manual valve, meaning that the cold jet flow remains constant throughout the run. As mentioned above, the Optimode™ modulation control unit offers a flexible programming of the cold jet flow, e.g. high flows at the beginning and low flows at the end of the GC temperature program.[50] Besides saving liquid nitrogen, this also may help to avoid peak tailing due to excessively high cold jet flows.[48]
- The length of the delay loop has an influence on modulation efficiency. While the analytes in a modulated fraction should be able to reach the second cold spot within one modulation period, their path should be long enough to prevent breakthrough within the hot jet pulse time. Based on the linear carrier gas velocity in the second dimension, the appropriate loop length can be calculated. Thus, depending on column dimensions, optimal loop length may vary between 0.6 and 1.6 m,[48, 50, 58] with about 1 m being probably a good compromise for most applications.

5.4.4. Temperature Program

Concerning the temperature program, many of the recommendations provided in section 5.3.3 are also valid for GC×GC. The main difference is that the temperature ramp should be much shallower than in one-dimensional GC, traditionally about 2-3°C/min[39, 41, 42, 69] and maximally 5-6°C/min,[23, 28, 34, 38, 51] depending on column dimensions or the complexity of the respective chromatogram region. The main reason is that shallow ramps are needed to make first-dimensional peaks broader than usual in order to meet the 'modulation criterion',[41] While it is tempting to use faster temperature programs also in GC×GC-MS metabolomic applications to reduce run time,[52, 67, 68] O'Hagan et al. observed that the number of detected peaks increases substantially if shallower oven ramps are used (see Fig. 5 in [23]). This indicates the importance of the chromatographic separation despite cutting-edge MS detection technologies and advanced data processing algorithms. It also makes clear that longer run times are the price to pay for the improved metabolome coverage as provided by GC×GC. However, as data processing, data evaluation and statistics are often the major bottlenecks in metabolomic workflows, longer run times appear, at the end, acceptable.

Generally, the slope of the GC temperature program is a means to control the t_R of the analytes; in case of GC×GC, this also applies for the second dimension. More precisely, the 2D retention of the analytes decreases if the slope of the ramp increases. Practically, to some extent, this effect can be used for fine-tuning of the 2D separation, to correct slight wrap-around or to reduce the modulation period.

5.5. Injection of LC Samples

5.5.1. Sample Injection Volume

The injection process must be relatively pulse-free to protect the column and the swept volume of the device needs to be minimal to reduce possible band spreading. Furthermore, to ensure high-throughput of samples, the injection cycle needs to be fast. In LC-MS applications, the sample injection volume can vary between 1 and 100 µL, but in ultra-high performance liquid chromatography (UHPLC), the ID of columns is limited and therefore, low injection volumes are necessary. Indeed, the sample volume injected must be proportional to the square of the ID of the column.[70] Typically, an injection volume between 1 and 10 µL is applied in metabolomic studies. As opposed to reverse-phase liquid chromatography (RP-LC), hydrophilic interaction liquid chromatography (HILIC) has low injection volume capability that can result in decreased sensitivity and resolution by generating wider peaks [71]. However, HILIC offers the possibility to inject the sample dissolved in organic solvents, which are preferable due to lower viscosity leading to lower back pressure.[71] Furthermore, organic solvents are weak eluents under HILIC conditions, so that the organic supernatant of extracted samples does not require evaporation and reconstitution.[72] Controlling the sample injection volume is very important in comprehensive metabolome profiling, since some metabolite concentrations are below the detection limit and others are highly abundant and lead to saturated signals. Also, over-injecting can result in sample carry-over from one injection to another.[73] Zhou and Li[73] performed a metabolomic analysis on $^{13}C_2$- and $^{12}C_2$-dansyl labelled urine using different injection volumes (2-12 µL). As a result, the peak pair number increased as the injection volume increased and stagnated more or less when the volume was 8 µL or higher, suggesting that the ionization process had become saturated.[73]

5.5.2. Full and Partial Loop Injection

Sample injection volume is also dependent on the loop volume used and different loop injections modes can be applied, i.e. full and partial loop injection. When using full loop mode, the maximum possible injection volume is determined by loop volume. However, when injecting a liquid through a capillary, parabolic flow profile may present, whereby the velocity of the liquid in the center is much higher than the liquid close to the capillary wall resulting in sample loss. Therefore, the sample loop must be 3-5 times overfilled as general rule when injecting in full loop mode. For example, when using a 5 µL loop, a sample volume between 15 and 25 µL must be flushed in the loop before injection.[74] In partially loop mode, a specific aliquot of the available loop volume is injected, whereby the injection volume does not exceed 50–75 % of the loop volume.[70] Thus, only 1-3 µL should be injected when using a 5 µL loop.[74] The use of partially loop mode is particularly interesting when a lower sample amount is available.[70]

5.6. Liquid Chromatography

Coupling chromatography to mass spectrometry offers an excellent solution to complex mixture analyses and has been extensively used in metabolomics. Chromatographic separation of metabolites prior to MS analyses has several advantages: i) reduces matrix effects and ionization

suppression, ii) separates isomers, iii) provides additional and orthogonal data (i.e. t_R/factor/index) valuable for metabolite annotation, and iv) allows for more accurate quantification of individual metabolites.

Nowadays, the development of UHPLC technology together with columns packed with sub-2 µm particles resulted in increased sensitivity, reduced matrix effects, and improved resolution.[75-78] The first generation of systems was limited to a certain flow rate to avoid too high back pressures. Nevertheless, recent evolutions in the technique enabled to work at higher flow rates and increased analysis speed. However, in most untargeted metabolomic studies, UHPLC is mainly applied to improve metabolite coverage by using relatively long gradient programs.[79] As suggested by Kohler and Giera,[79] untargeted analysis would rather benefit from UHPLC high-throughput analysis, for example by performing an initial fast screening followed by quantification and/or identification of relevant metabolites.

5.6.1. Columns Used in LC-MS

In recent years, a large number of different types of columns have been introduced in the -omics field, including reversed phase (RP), normal phase (NP) and HILIC columns, each relying on different mechanisms for separation and retention of metabolites. Currently, it is not possible to use only one of these techniques to profile the whole metabolome and therefore complementary techniques are frequently combined to increase metabolite coverage. In 2007, core-shell particles, made of a solid core surrounded by silica porous particles, were introduced and led to better efficiency without the need for high pressure due to better eddy diffusion, mass-transfer resistance and a decrease in longitudinal diffusion. Nevertheless, this technology is mostly limited to lipidomics applications and its lower use in metabolomic studies could possibly be explained by its lower performance observed with some columns induced by a smaller interaction area.[79] For example, Kloos et al.[80] suggested that the core-shell pentafluorophenyl (PFP)-RP-LC material does not perform as well as its porous equivalent, especially in the analysis of hydrophilic analytes.[80]

RP-LC

RP-LC has been extensively used in LC-MS based untargeted -omic profiling due to its high versatility, stability, and large metabolite coverage.[81] It is considered to be the most suitable analytical tool for ESI-MS based analysis of complex samples.[82] RP-LC relies on hydrophobic interactions and is suitable to mainly separate and retain nonpolar and medium polar analytes. However, RP-LC, especially with C8 and C18 stationary phases, is not the best method to separate polar compounds (such as sugars and amino acids), since they generally elute simultaneously near the column void volume and thus t_R cannot give valuable information about their identification.[83] However, improved retention of polar compounds can be achieved with recently developed commercially RP columns that incorporate an amide or carbamate into the alkyl ligand, such as the T3 (Waters, Milford, MA, USA), SB-Aq (Agilent, Santa Clara, CA, USA), and Synergi Hydro-RP (Phenomonex, Torrance, CA, USA). Nevertheless, these columns can only minimally retain highly polar compounds, including sugars, carboxylic acids, and most amino acids.[84] Contrepois et al.[85] compared five C18-bonded silica RP-LC columns (Hypersil GOLD, Hypersil GOLD aq., BEH C18, Kinetex, Zorbax SB) for global metabolite profiling of plasma and urine and concluded that Hypersil GOLD and Zorbax SB aq columns gave the best results.[85] In the case of lipidomics, i.e. analysis of nonpolar compounds, RP-LC is a valuable analytical tool in the separation of lipid species. Nevertheless, under commonly used RP-LC conditions phosphatidic acids and phosphatidylserines, both phospholipid classes, tend to elute as extensively broad peaks leading to the use of NP-LC in the analysis of these phospholipid classes.[86] RP-LC columns are mostly operated at a flow rate of 0.1-0.5 mL/min and maintained at temperatures between 40-55°C.[86]

Aqueous NP-LC

The use of aqueous NP-LC enables to retain longer more polar molecules on the stationary phase, whereby direct interaction with the latter one allows both hydrophilic and hydrophobic retention.[87] This method has already proven its complementarity to RP-LC in urinary metabolomics, whereby retention of creatine and creatinine could only be observed with aqueous NP-LC and not with RP-LC.[84]

Non-Aqueous NP-LC

As mentioned before, non-aqueous NP-LC is mainly used in the separation of lipid classes and is based on the physical properties of the lipid head groups resulting in co-elution of lipids belonging to the same class or subclass.[88] For NP-LC columns, a flow rate between 0.1 and 0.5 mL/min and a temperature of 20-35°C is frequently applied.[86]

HILIC

Another LC method that offers the opportunity to separate polar compounds is HILIC. This method represents a complementary or alternative method to RP-LC and NP-LC, respectively, achieving separation and retention of molecules by hydrophilic partitioning between an organic solvent-rich mobile phase and an aqueous layer that is formed on the stationary phase.[83] The formation of the water-enriched layer, which is absent in NP-LC, is made possible by the 5-40 % water content of the mobile phase.[79, 86] The stationary phases of HILIC columns can be classified into various types based on the charge characteristics of the functional groups, i.e. neutral (silica derivatized with mostly an uncharged amide group), anionic (mostly bare silica), cationic (silica derivatized with mostly an aminopropyl positively charged group), and zwitterionic phases (silica derivatized with mostly sulfobetaine bearing a positive and negative charge).[85, 89] Contrepois et al.[85] demonstrated that the zwitterionic ZIC HILIC column gave better results for urinary and plasma metabolite profiling in comparison to the other columns BEH amide (uncharged stationary phase), BEH HILIC (anionic stationary phase), Hypersil GOLD HILIC (uncharged stationary phase), and Synchronis HILIC (zwitterionic stationary phase).[85] Compared to RP-LC, HILIC requires longer equilibration time in gradient elution programs that is achieved by multiple injecting QC samples before analysis of the study samples. Additionally, HILIC can suffer from poor t_R repeatability and low peaks when important factors, such as injection volume and use of buffers, are not considered carefully.[79, 90] HILIC columns are operated with a flow rate of 0.1 - 1 mL/min and temperature of 25-40°C.[86]

5.6.2. Two Dimensional (2D)-LC Technology

As discussed before, it is impossible to cover the whole metabolome by the use of one specific LC-MS method. Therefore, different approaches are gaining interest in untargeted metabolomics such as the serial combining of columns, column switching, and two dimensional (2D)-LC technology.[91] In the first method, different column chemistries and lengths can be coupled based on their orthogonality and compatibility in mobile phases. For example, the literature has already demonstrated the usefulness of the combination of RP-LC with HILIC for the detection of both polar and nonpolar compounds in a heterogeneous sample.[92, 93] In column switching applications, a six-port or ten-port valve is connected to two separate pumps and columns and allows analyzing samples on different column chemistries in one analytical run. As opposed to serial combining of columns, mobile phases used for the different columns do not need to be compatible. However, in this method, two injections are required for one full analysis.[91] Li et al.[94] used this technology for the analysis of metabolites and lipids in human plasma.[94] Finally, 2D-LC analysis can be divided in two different categories, i.e. complementary and heart-cutting 2D-LC. In the first method, all the fractions that have run through a first column are collected in aliquots and injected through a second column and both dimensions have very short gradients. In the second method, only a selection of fractions (metabolites of interest) are subjected to the second dimension, which has longer gradients than the first dimension. The transfer of fractions can be performed off-line or on-line (automated). The main advantage of 2D-LC is the combination of two completely different LC separation platforms, whereby incompatibility of mobile phases is not an issue.[91, 95]

5.6.3. Supercritical Fluid Chromatography (SFC)

Supercritical fluid chromatography (SFC) is another technique that is mainly applied in the field of lipidomics, due to efficient lipid class separation, and its use has significantly increased over the last few years. The method relies on a liquid mobile phase (typically carbon dioxide added with 2-30 % methanol or ethanol), whereby the fluid is maintained beyond its critical point resulting in similar diffusity of a gas and similar solubility of a liquid. These characteristics enables the use of higher flow rates resulting in shorter analysis time without a significant increase in back pressure.[79]

5.6.4. High-Temperature Liquid Chromatography (HTLC)

High-temperature liquid chromatography (HTLC) could be beneficial in untargeted analyses, since high temperature results in a decrease of viscosity of the mobile phase, back pressure, and the amount of organic solvent needed. Furthermore, it allows the use of high flow rates by the enhancement of mass transfer.[79] However, there is reluctance to use this technique, since the behavior of thermally unstable metabolites needs to be further investigated and since only a small number of columns are currently stable at high temperatures.[96, 97] Moreover, plasma and urine samples gave contradictory results when an HTLC-UHPLC method was compared to a conventional LC method in a study of Gika et al.[98]

5.6.5. Mobile Phases

Gradient Elution

In untargeted metabolomic studies gradient elution is preferred to isocratic elution, since the metabolome comprises a wide range of polarities. In gradient solvent systems, the mobile phase is not constant as the analysis progresses, since the eluotropic strength of the solvent is increased. The choice of solvents used as mobile phases is dependent on column chemistry. For example, in RP-LC the starting conditions of the mobile phase are predominantly aqueous (99-100%) to ensure retention of medium polar and nonpolar molecules. By increasing the organic solvent content (most often ACN or MeOH), the analytes are eluted from the column and introduced into the MS system. Recently, the use of MeOH as organic phase for RP-LC resulted in higher peak intensities, particularly in negative ionization mode (increase up to 25-fold) as opposed to ACN in a plasma metabolome study. Additionally, peak broadening effects, which are expected with the use of MeOH, were negligible in respect to the gain in improvement of peak intensity.[99] In HILIC, the gradient usually starts with 95% ACN whereby the aqueous content in the mobile phase gradually increases up to 90%.[71, 72] ACN is the strongly preferred organic solvent in HILIC applications since it gives better sample retention and peak shapes as opposed to MeOH.[72]

Buffers

Buffers are commonly added to mobile phases, because they can positively influence the retention, separation, and ionization of molecules. Indeed, buffers determine the mobile phase pH and the selectivity of stationary phase.[72, 81] The use of additives is especially of interest when compounds contain ionizable characteristics such as amino or carboxylic groups.[100] Mainly formic acid, acetic acid, ammonium acetate, ammonium formatate, ammoniumhydroxide, and trifluoracetic acid (TFA) are added to mobile phases.[86, 100] Especially, ammonium formate and ammonium acetate are popular additives in HILIC applications.[100, 101] Formic acid is frequently used in RP-LC to reduce pH and suppress ionization of weak organic acids for urine metabolome studies.[3] These acids or bases can have various concentration-dependent effects on ESI so special care must be taken. Indeed, the formation of particular charged adducts such as $[M+NH_4]^+$, $[M+FA]^-$, and $[M+AcO^-]^-$ can suppress peak height and sensitivity and increase spectral complexity.[100, 102] To illustrate this, Ivanisevic et al.[81] tested different concentrations (5, 10, 20, 50 mM) of ammonium acetate and ammonium hydroxide and concluded that the first three concentrations enhanced the ESI sensitivity in HILIC-MS, whilst the use of 50 mM led to a significant decrease in sensitivity possibly due to ion suppression. Moreover, in this study different pH conditions (4.0, 6.9 and 9.8) were tested and pH 9.8 resulted in the highest number of resolved molecules in both positive and negative ionization mode. These effects can possibly be explained by weaker electrostatic interactions between the aminopropyl stationary phase and negatively charged compounds and a better deprotonation of acidic analytes resulting in improved MS sensitivity.[81] In lipidomic studies, the concentration range of ammonium acetate, formate and hydroxide is usually between 5–20 mM, and formic acid or acetic acid is frequently added in concentrations between 0.05-1%.[86] Ogiso et al.[103] added 5 µM of phosphoric acid

into the mobile phase to reduce peak tailing of phosphatidic acid (PA) by suppressing the unidentified interaction between PA and the column material.[103] Tailing of peaks is also a phenomenon that is more common in HILIC than in RP-LC, since HILIC columns are more easily overloaded. Addition of TFA can improve peak shapes of basic compounds on silica-based columns.[72]

Post Column Modifiers

Post column modifiers, i.e. addition of organic solvents to the mobile phase at the LC-MS interface, can enhance ESI sensitivity, but have unfortunately not been explored so far in untargeted metabolomics. Koch et al.[100] demonstrated that post-column infusion of 2-(2-methoxyethoxy)ethanol (2-MEE) in a HILIC application improved ESI efficiency for especially small and polar compounds in the negative ionization mode. Additionally, Tulipani et al.[99] showed that post-column enrichment of the organic phase with MeOH in RP-LC also enhanced signal intensity, this by reducing the aqueous fraction that reaches the MS detector.

To conclude, it is advisable to test a wide range of columns and gradients to assess which combination enables to detect the highest number of features in a particular experimental set-up, since different column chemistries and solvents interact with each other in a different way and the choice is dependent of the molecules of interest. This is further illustrated by studies for global metabolite profiling of plasma and urine published in 2013, which used different columns and chromatographic conditions and reached no consensus on which method produced the best reproducibility and metabolome coverage.[85] Also, special attention must be paid to not shorten the analysis time too much, thus a reasonable compromise between speed of analysis time and metabolome coverage must be made.[3] In most untargeted metabolomic studies of different biofluids gradient elution time is typically 15-30 min on 100 mm x 2.1 mm I.D.[15, 76, 104-106] Cajka et al.[86] reviewed lipidomic studies and came to the conclusion that in 65% and in 45% of the cases reported analysis times were less than 30 min and less than 20 min, respectively.

References

[1] L. Sumner, A. Amberg, D. Barrett, M. Beale, R. Beger, C. Daykin, T. M. Fan, O. Fiehn, R. Goodacre, J. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. Lane, J. Lindon, P. Marriott, A. Nicholls, M. Reily, J. Thaden, M. Viant, *Metabolomics* **2007**, 3, 211.

[2] J. M. Baker, N. D. Hawkins, J. L. Ward, A. Lovegrove, J. A. Napier, P. R. Shewry, M. H. Beale, *Plant Biotechnol. J.* **2006**, 4, 381.

[3] E. J. Want, I. D. Wilson, H. Gika, G. Theodoridis, R. S. Plumb, J. Shockcor, E. Holmes, J. K. Nicholson, *Nat. Protoc.* **2010**, 5, 1005.

[4] M. Berg, M. Vanaerschot, A. Jankevics, B. Cuypers, R. Breitling, J. C. Dujardin, *Comput. Struct. Biotechnol. J.* **2013**, 4, 201301002.

[5] F. M. Van Der Kloet, Ivana Bobeldijk, Elwin R. Verheij, R. H. Jellema, *J. Proteome Res.* **2009**, 8, 1272.

[6] U. Roessner, C. Wagner, J. Kopka, R. N. Trethewey, L. Willmitzer, *Plant J.* **2000**, 23, 131.

[7] A. Quero, C. Jousse, M. Lequart-Pillon, E. Gontier, X. Guillot, B. Courtois, J. Courtois, C. Pau-Roblot, *J. Chromatogr. B* **2014**, 970, 36.

[8] G. Moros, A. C. Chatziioannou, H. G. Gika, N. Raikos, G. Theodoridis, *Bioanalysis* **2017**, 9, 53.

[9] P. Begley, S. Francis-McIntyre, W. B. Dunn, D. I. Broadhurst, A. Halsall, A. Tseng, J. Knowles, R. Goodacre, D. B. Kell, *Anal. Chem.* **2009**, 81, 7038.

[10] E. Zarate, V. Boyle, U. Rupprecht, S. Green, S. G. Villas-Boas, P. Baker, F. R. Pinu, *Metabolites* **2017**, 7.

[11] H. G. Gika, G. A. Theodoridis, J. E. Wingate, I. D. Wilson, *J. Proteome Res.* **2007**, 6, 3291.

[12] W. B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J. D. Knowles, A. Halsall, J. N. Haselden, A. W. Nicholls, I. D. Wilson, D. B. Kell, R. Goodacre, *Nat. Protoc.* **2011**, 6, 1060.

[13] M. Koek, F. Van Der Kloet, R. Kleemann, T. Kooistra, E. Verheij, T. Hankemeier, *Metabolomics* **2011**, 7, 1.

[14] M. A. Kamleh, T. M. Ebbels, K. Spagou, P. Masson, E. J. Want, *Anal. Chem.* **2012**, 84, 2670.

[15] J. Vanden Bussche, M. Marzorati, D. Laukens, L. Vanhaecke, *Anal. Chem.* **2015**, 87, 10927.

[16] K. Grob, *Split and Splitless Injection for Quantitative Gas Chromatography: Concepts, Processes, Practical Guidelines, Sources of Error*, Wiley-VCH, Weinheim, Germany 2001.

[17] S. Lacorte, A. R. Fernández-Alba, in: Fernández-Alba, A. R. (Ed.), *Compr. Anal. Chem.* 2005, 43, 235.

[18] J. Gullberg, P. Jonsson, A. Nordström, M. Sjöström, T. Moritz, *Anal. Biochem.* **2004**, 331, 283.

[19] C. Constantinou, P. K. Chrysanthopoulos, M. Margarity, M. I. Klapa, *J. Proteome Res.* **2011**, 10, 869.

[20] A. Jiye, J. Trygg, J. Gullberg, A. I. Johansson, P. Jonsson, H. Antti, S. L. Marklund, T. Moritz, *Anal. Chem.* **2005**, 77, 8086.

[21] A. Jiye, Q. Huang, G. Wang, W. Zha, B. Yan, H. Ren, S. Gu, Y. Zhang, Q. Zhang, F. Shao, L. Sheng, J. Sun, *Anal. Biochem.* **2008**, 379, 20.

[22] L. H. Münger, A. Trimigno, G. Picone, C. Freiburghaus, G. Pimentel, K. J. Burton, F. P. Pralong, N. Vionnet, F. Capozzi, R. Badertscher, G. Vergeres, *J. Proteome Res.* **2017**, 16, 3321.

[23] S. O'Hagan, W. B. Dunn, J. D. Knowles, D. Broadhurst, R. Williams, J. J. Ashworth, M. Cameron, D. B. Kell, *Anal. Chem.* **2007**, 79, 464.

[24] K. K. Pasikanti, P. C. Ho, E. C. Chan, *Rapid Commun. Mass Spectrom.* **2008**, 22, 2984.

[25] X. Li, Z. Xu, X. Lu, X. Yang, P. Yin, H. Kong, Y. Yu, G. Xu, *Anal. Chim. Acta* **2009**, 633, 257.

[26] N. Psychogios, D. D. Hau, J. Peng, A. C. Guo, R. Mandal, S. Bouatra, I. Sinelnikov, R. Krishnamurthy, R. Eisner, B. Gautam, N. Young, J. Xia, C. Knox, E. Dong, P. Huang, Z. Hollander, T. L. Pedersen, S. R. Smith, F. Bamforth, R. Greiner, B. McManus, J. W. Newman, T. Goodfriend, D. S. Wishart, *PloS One* **2011**, 6, 16957.

[27] S. Bouatra, F. Aziat, R. Mandal, A. C. Guo, M. R. Wilson, C. Knox, T. C. Bjorndahl, R. Krishnamurthy, F. Saleem, P. Liu, Z. T. Dame, J. Poelzer, J. Huynh, F. S. Yallou, N. Psychogios, E. Dong, R. Bogumil, C. Roehring, D. S. Wishart, *PloS One* **2013**, 8, 73076.

[28] C. H. Weinert, B. Egert, S. E. Kulling, *J. Chromatogr. A* **2015**, 1405, 156.

[29] J. W. Allwood, A. Erban, S. De Koning, W. B. Dunn, A. Luedemann, A. Lommen, L. Kay, R. Loscher, J. Kopka, R. Goodacre, *Metabolomics* **2009**, 5, 479.

[30] E. C. Y. Chan, K. K. Pasikanti, J. K. Nicholson, *Nat. Protoc.* **2011**, 6, 1483.

[31] H.-J. Hübschmann, *Fundamentals. In: Handbook of GC/MS: Fundamentals and Applications*, Wiley-VCH, Weinheim, Germany 2008.

[32] W. Zou, J. She, V. V. Tolstikov, *Metabolites* **2013**, 3, 787.

[33] W. S. Law, P. Y. Huang, E. S. Ong, C. N. Ong, S. F. Y. Li, K. K. Pasikanti, E. C. Y. Chan, *Rapid Commun. Mass Spectrom.* **2008**, 22, 2436.

[34] C. H. Weinert, M. T. Empl, R. Kruger, L. Frommherz, B. Egert, P. Steinberg, S. E. Kulling, *Mol. Nutr. Food Res.* **2017**, 61, 1.

[35] S. M. M. Pedersen, C. Nebel, N. C. Nielsen, H. J. Andersen, J. Olsson, M. Simren, L. Ohman, U. Svensson, H. C. Bertram, A. Malmendal, *Eur. Food Res. Technol.* **2011**, 233, 1013.

[36] J. C. Giddings, *J. High Res. Chromatogr.* **1987**, 10, 319.

[37] J. Dallüge, R. J. J. Vreuls, J. Beens, U. a. T. Brinkman, *J. Sep. Sci.* **2002**, 25, 201.
[38] J. Dallüge, J. Beens, U. a. T. Brinkman, *J. Chromatogr. A* **2003**, 1000, 69.
[39] M. Adahchour, J. Beens, R. J. J. Vreuls, U. a. T. Brinkman, TrAC *Trends Anal. Chem.* **2006**, 25, 438.
[40] M. Adahchour, J. Beens, R. J. J. Vreuls, U. a. T. Brinkman, TrAC *Trends Anal. Chem.* **2006**, 25, 540.
[41] M. Adahchour, J. Beens, U. a. T. Brinkman, *J. Chromatogr. A* **2008**, 1186, 67.
[42] L. Mondello, P. Q. Tranchida, P. Dugo, G. Dugo, *Mass Spectrom. Rev.* **2008**, 27, 101.
[43] P. J. Marriott, S.-T. Chin, B. Maikhunthod, H.-G. Schmarr, S. Bieri, *TrAC Trends Anal. Chem.* **2012**, 34, 1.
[44] G. Semard, M. Adahchour, J.-F. Focant, in: Lourdes, R. (Ed.), *Comprehensive Two Dimensional Gas Chromatography*, Elsevier, Amsterdam 2009, pp. 15-48.
[45] M. F. Almstetter, P. J. Oefner, K. Dettmer, *Anal. Bioanal. Chem.* **2012**, 402, 1993.
[46] J. Krupčík, R. Gorovenko, I. Špánik, P. Sandra, D. W. Armstrong, *J. Chromatogr. A* **2013**, 1280, 104.
[47] M. Edwards, A. Mostafa, T. Górecki, *Anal. Bioanal. Chem.* **2011**, 401, 2335.
[48] P. Q. Tranchida, G. Purcaro, P. Dugo, L. Mondello, G. Purcaro, *TrAC Trends Anal. Chem.* **2011**, 30, 1437.
[49] P. Q. Tranchida, F. A. Franchina, P. Dugo, L. Mondello, *Mass Spectrometry Reviews* **2016**, 35, 524.
[50] D. Bressanello, E. Liberto, M. Collino, S. E. Reichenbach, E. Benetti, F. Chiazza, C. Bicchi, C. Cordero, *J. Chromatogr. A* **2014**, 1361, 265.
[51] J. H. Winnike, X. Wei, K. J. Knagge, S. D. Colman, S. G. Gregory, X. Zhang, *J. Proteome Res.* **2015**, 14, 1810.
[52] M. F. Almstetter, I. J. Appel, M. A. Gruber, C. Lottaz, B. Timischl, R. Spang, K. Dettmer, P. J. Oefner, *Anal. Chem.* **2009**, 81, 5731.
[53] J. Harynuk, T. Górecki, J. D. Zeeuw, *J. Chromatogr. A* **2005**, 1071, 21.
[54] M. M. Koek, B. Muilwijk, L. L. P. Van Stee, T. Hankemeier, *J. Chromatogr. A* **2008**, 1186, 420.
[55] J. Beens, H.-G. Janssen, M. Adahchour, U. a. T. Brinkman, *J. Chromatogr. A* **2005**, 1086, 141.
[56] P. Q. Tranchida, A. Casilli, P. Dugo, G. Dugo, L. Mondello, *Anal. Chem.* **2007**, 79, 2266.
[57] P. Q. Tranchida, G. Purcaro, L. Conte, P. Dugo, G. Dugo, L. Mondello, *J. Chromatogr. A* **2009**, 1216, 7301.
[58] L. Nicolotti, C. Cordero, D. Bressanello, C. Cagliero, E. Liberto, F. Magagna, P. Rubiolo, B. Sgorbini, C. Bicchi, *J. Chromatogr. A* **2014**, 1360, 264.
[59] P. Q. Tranchida, G. Purcaro, L. Conte, P. Dugo, G. Dugo, L. Mondello, *Anal. Chem.* **2009**, 81, 8529.
[60] E. B. Ledford, C. Billesbach, *J. High Res. Chromatogr.* **2000**, 23, 202.
[61] J.-F. Focant, A. Sjödin, D. G. Patterson, *J. Chromatogr. A* **2003**, 1019, 143.
[62] E. B. Ledford, C. Billesbach, J. Termaat, *Pittcon*, New Orleans, USA **2002**.
[63] P. Q. Tranchida, F. A. Franchina, P. Dugo, L. Mondello, *J. Chromatogr. A* **2014**, 1359, 271.
[64] F. A. Franchina, M. Maimone, P. Q. Tranchida, L. Mondello, *J. Chromatogr. A* **2016**, 1441, 134.
[65] B. Egert, C. H. Weinert, S. E. Kulling, *J. Chromatogr. A* **2015**, 1405, 168.
[66] W. Welthagen, R. Shellie, J. Spranger, M. Ristow, R. Zimmermann, O. Fiehn, *Metabolomics* **2005**, 1, 65.
[67] K. K. Pasikanti, J. Norasmara, S. Cai, R. Mahendran, K. Esuvaranathan, P. C. Ho, E. C. Y. Chan, *Anal. Bioanal. Chem.* **2010**, 398, 1285.
[68] M. Hartonen, I. Mattila, A.-L. Ruskeepää, M. Orešič, T. Hyötyläinen, *J. Chromatogr. A* **2013**, 1293, 142.
[69] P. Q. Tranchida, R. A. Shellie, G. Purcaro, L. S. Conte, P. Dugo, G. Dugo, L. Mondello, *J. Chromatogr. Sci.* **2010**, 48, 262.
[70] C. E. Nazario, M. R. Silva, M. S. Franco, F. M. Lancas, *J. Chromatogr. A* **2015**, 1421, 18.
[71] A. S. Lopes, E. C. Cruz, A. Sussulini, A. Klassen, *Adv. Exp. Med. Biol.* **2017**, 965, 77.
[72] P. Jandera, *Anal. Chim. Acta* **2011**, 692, 1.
[73] R. Zhou, L. Li, *J. Proteomics* **2015**, 118, 130.
[74] L. Mondello, *Anal. Bioanal. Chem.* **2017**, 409, 4481.
[75] D. Guillarme, J. Ruta, S. Rudaz, J. L. Veuthey, *Anal. Bioanal. Chem.* **2010**, 397, 1069.
[76] H. G. Gika, G. A. Theodoridis, R. S. Plumb, I. D. Wilson, *J. Pharmaceut. Biomed.* **2014**, 87, 12.
[77] S. Fekete, J. Schappler, J.-L. Veuthey, D. Guillarme, *TrAC Trends Anal. Chem.* **2014**, 63, 2.
[78] S. Fekete, J. L. Veuthey, D. Guillarme, *J. Chromatogr. A* **2015**, 1408, 1.
[79] I. Kohler, M. Giera, *J. Sep. Sci.* **2017**, 40, 93.
[80] D. P. Kloos, H. Lingeman, W. M. Niessen, A. M. Deelder, M. Giera, O. A. Mayboroda, *J. Chromatogr. B* **2013**, 927, 90.
[81] J. Ivanisevic, Z. J. Zhu, L. Plate, R. Tautenhahn, S. Chen, P. J. O'Brien, C. H. Johnson, M. A. Marletta, G. J. Patti, G. Siuzdak, *Anal. Chem.* **2013**, 85, 6876.
[82] G. A. Theodoridis, H. G. Gika, E. J. Want, I. D. Wilson, *Anal. Chim. Acta* **2012**, 711, 7.
[83] T. Zhang, D. J. Creek, M. P. Barrett, G. Blackburn, D. G. Watson, *Anal. Chem.* **2012**, 84, 1994.
[84] D. L. Callahan, D. De Souza, A. Bacic, U. Roessner, *J. Sep. Sci.* **2009**, 32, 2273.
[85] K. Contrepois, L. Jiang, M. Snyder, *Mol. Cell. Proteomics* **2015**, 14, 1684.
[86] T. Cajka, O. Fiehn, TrAC *Trends Anal. Chem.* **2014**, 61, 192.
[87] J. J. Pesek, M. T. Matyska, S. Larrabee, *J. Sep. Sci.* **2007**, 30, 637.
[88] X. Gao, Q. Zhang, D. Meng, G. Isaac, R. Zhao, T. L. Fillmore, R. K. Chu, J. Zhou, K. Tang, Z. Hu, R. J. Moore, R. D. Smith, M. G. Katze, T. O. Metz, *Anal. Bioanal. Chem.* **2012**, 402, 2923.
[89] K. Inoue, H. Tsuchiya, T. Takayama, H. Akatsu, Y. Hashizume, T. Yamamoto, N. Matsukawa, T. Toyo'oka, *J. Chromatogr. B* **2015**, 974, 24.
[90] T. Zhang, D. G. Watson, *J. Chromatogr. B* **2016**, 1022, 199.
[91] J. Haggarty, K. E. Burgess, *Curr. Opin. Biotech.* **2017**, 43, 77.
[92] K. R. Chalcraft, B. E. McCarry, *J. Sep. Sci.* **2013**, 36, 3478.
[93] J. Haggarty, M. Oppermann, M. J. Dalby, R. J. Burchmore, K. Cook, S. Weidt, K. E. Burgess, *Metabolomics* **2015**, 11, 1465.
[94] Y. Li, Z. Zhang, X. Liu, A. Li, Z. Hou, Y. Wang, Y. Zhang, *J. Chromatogr. A* **2015**, 1409, 277.
[95] P. Yin, G. Xu, *J. Chromatogr. A* **2014**, 1374, 1.
[96] D. Guillarme, J. Schappler, S. Rudaz, J.-L. Veuthey, *TrAC Trends Anal. Chem.* **2010**, 29, 15.
[97] P. D. Rainville, G. Theodoridis, R. S. Plumb, I. D. Wilson, *TrAC Trends Anal. Chem.* **2014**, 61, 181.
[98] H. G. Gika, G. A. Theodoridis, I. D. Wilson, *J. Sep. Sci.* **2008**, 31, 1598.
[99] S. Tulipani, X. Mora-Cubillos, O. Jauregui, R. Llorach, E. Garcia-Fuentes, F. J. Tinahones, C. Andres-Lacueva, *Anal. Chem.* **2015**, 87, 2639.
[100] W. Koch, S. Forcisi, R. Lehmann, P. Schmitt-Kopplin, *J. Chromatogr. A* **2014**, 1361, 209.
[101] Y. Guo, S. Gaiki, *J. Chromatogr. A* **2005**, 1074, 71.
[102] N. Huang, M. M. Siegel, G. H. Kruppa, F. H. Laukien, *J. Am. Soc. Mass Spectr.* **1999**, 10, 1166.
[103] H. Ogiso, T. Suzuki, R. Taguchi, *Anal. Biochem.* **2008**, 375, 124.
[104] A. M. Evans, C. D. Dehaven, T. Barrett, M. Mitchell, E. Milgram, *Anal. Chem.* **2009**, 81, 6656.
[105] W. B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J. D. Knowles, A. Halsall, J. N. Haselden, A. W. Nicholls, I. D. Wilson, D. B. Kell, R. Goodacre, C. Human Serum Metabolome, *Nat. Protoc.* **2011**, 6, 1060.
[106] T. Fuhrer, N. Zamboni, *Curr. Opin. Biotech.* **2015**, 31, 73.