



HAL
open science

Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: A review.

Aurélie Roux, Dominique Lison, Christophe Junot, Jean-François Heillier

► To cite this version:

Aurélie Roux, Dominique Lison, Christophe Junot, Jean-François Heillier. Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: A review.. *Clinical Biochemistry*, 2011, 44 (1), pp.119-135. hal-00641535

HAL Id: hal-00641535

<https://hal.science/hal-00641535>

Submitted on 16 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: a review

Aurélie Roux^a, Dominique Lison^b, Christophe Junot^{a*} and Jean-François Heilier^{a, b, c}

^a Service de Pharmacologie et d'Immunoanalyse, DSV/iBiTec-S, CEA/Saclay, 91191 Gif-sur-Yvette cedex, France.

^b Université catholique de Louvain, Louvain centre for Toxicology and Applied Pharmacology (LTAP), 1200 - Brussels, Belgique.

^c Institut National de Recherche et de Sécurité. Laboratoire de Surveillance Biologique de l'Exposition aux Substances Inorganiques. 54519 Vandœuvre-lès-Nancy.

*Author for correspondence. Phone: 33-1-69-08-43-66. Fax: 33-1-1-69-08-59-07.

christophe.junot@cea.fr

Abstract

The metabolome is the set of small molecular mass organic compounds found in a given biological media. It includes all organic substances naturally occurring from the metabolism of the studied living organism, except biological polymers, but also xenobiotics and their biotransformation products. The metabolic fingerprints of biofluids obtained by mass spectrometry (MS) or nuclear magnetic resonance (NMR)-based methods contain a few hundreds to thousands of signals related to both genetic and environmental contributions. Metabolomics, which refers to the untargeted quantitative or semi-quantitative analysis of the metabolome, is a promising tool for biomarker discovery. Although proof-of-concept studies by metabolomics-based approaches in the field of toxicology and clinical chemistry have initially been performed using NMR, the use of liquid chromatography hyphenated to mass spectrometry (LC/MS) has increased over the recent years, providing complementary results to those obtained with other approaches. This paper reviews and comments the input of LC/MS in this field. We describe here the overall process of analysis, review some seminal papers in the field and discuss the perspectives of metabolomics for the biomonitoring of exposure and diagnosis of diseases.

Keywords: metabolomics, metabonomics, LC/MS, toxicology, biomarkers,

Abbreviations

MS: mass spectrometry; NMR: nuclear magnetic resonance; LC: liquid chromatography; MHz: MegaHertz; UPLC: Ultra performance liquid chromatography; HPLC: High performance liquid chromatography; TOF: time of flight; MS/MS: tandem mass spectrometry; API: atmospheric pressure ionisation; GSH: glutathione; NAPQI: N-acetyl-*p*-quinone-imine; APAP: *N*-acetyl-*p*-aminophenol; GC: gas chromatography; MMC: Methyl mercury chloride; RC: respiratory chain; RCD: respiratory chain disease; Cr: creatine; PCr : phosphocreatine; CRC: Colorectal cancer; FT-ICR: Fourier Transform Ion Cyclotron Resonance; Q-Trap: Quadrupole linear trap; Q-TOF: Quadrupole Time-of-flight; CID : collision induced dissociation; DNA: Deoxyribonucleic acid; RNA: Ribonucleic acid; ESI: Electrospray ionization; QC: Quality control; PCA: Principal Components Analysis; PLS: Partial least squares or projection to latent structures; PLS-DA: Partial least squares discriminant analysis; OPLS: Orthogonal partial least squares; ALT: Alanine transaminase; AST: Aspartate transaminase; BUN: Blood urea nitrogen; ATP: Adenosine triphosphate; PC: Phosphatidylcholine; PSA: Prostate-Specific Antigen; TCA cycle: Tricarboxylic acid cycle; D-AAO: D-amino-acid oxidase; CYP: Cytochrome P450; PPAR α : Peroxisome proliferator-activated receptor alpha; ANIT: R-naphthyl isothiocyanate; CCl₄: carbon tetrachloride; LDH: lactate dehydrogenase; GABA: γ -Aminobutyric acid ; H/D exchange : Hydrogen-deuterium exchange

Contents

Introduction	5
<i>What is a metabolite?</i>	5
<i>Metabolomics: a new approach for biomarker discovery</i>	6
<i>How to measure metabolites?</i>	9
LC/MS based metabolomics: a practical approach.....	10
<i>Design of the experiment</i>	11
<i>Sample preparation</i>	12
<i>Acquisition of metabolic fingerprints using LC/ESI-MS systems</i>	13
<i>Automatic detection of ions</i>	14
<i>Statistical analyses</i>	15
<i>Identification</i>	16
Selected applications in the field of toxicology	18
<i>LC/MS based metabolomics for toxicity biomarkers discovery</i>	20
<i>LC/MS-based metabolomics for the building of predictive models of toxicity</i>	22
<i>Toward mechanistic considerations</i>	24
Selected applications in the field of clinical chemistry.....	25
Conclusion and perspectives	30

Introduction

The metabolome is a set of small molecular mass organic compounds found in a given biological medium. Polymerized structures such as proteins and nucleic acids are excluded from the metabolome but small peptides such as the tripeptide glutathione are included. Molecules that constitute the metabolome are called metabolites.

What is a metabolite?

For some scientists, the concept of metabolite includes all the organic substances naturally occurring from the metabolism of a living organism and that do not directly come from gene expression. It should be stressed here that this definition could be applied as well to a microorganism, a human being or a plant. Two different kinds of metabolites can be distinguished based on their origin: endogenous and exogenous metabolites.

Endogenous metabolites could be classified as primary and secondary metabolites. The firsts have a broad distribution in living species and are directly involved in essential life processes such as growth, development and reproduction. This is for example the case for amino-acids or glycolysis intermediates. At the opposite, secondary metabolites are species-specific, have a restricted distribution and are synthesized for a particular biological function, as alkaloids for plants or hormones for mammals [1].

Exogenous metabolites represent the biotransformation or metabolism products of exogenous compounds, resulting from phase I (modification of the original molecule to introduce a functional group) and/or phase II (conjugation) enzymatic conversion [2]. In this particular context, Holmes et al [3] proposed the concept of xenometabolome which is a description of the xenobiotic metabolite profile of an individual exposed to environmental pollutants, drugs, or exogenous molecules coming from food/dietary components such as phytochemicals [4]. This concept expands the approach developed in the early nineties in the

field of molecular epidemiology [5;6], thanks to the technical advances in analytical chemistry.

In epidemiological studies, analyzing the xenometabolome could especially allow characterising environmental or occupational exposures to chemicals and contributes therefore to the determination of a metabolic phenotype. Crockford et al [7] demonstrated the potential of this approach by identifying metabolites of drugs such as acetaminophen or disopyramide by heterospectroscopy on data acquired with 600-MHz ^1H NMR and UPLC-TOF-MS^E on urines obtained from more than 80 patients.

Metabolomics: a new approach for biomarker discovery

Biochemists have long been doing metabolomics, just like the Bourgeois Gentilhomme was speaking prose without knowing it (Molière – Bourgeois Gentilhomme II. 4). It means that they suspected that patterns of biochemical substances could explain or describe inter-individual variation. Gates and Sweeley [8] mention that the concept of metabolic pattern was introduced by Williams [9;10] who used paper chromatography to compare the urines of 200,000 subjects including alcoholics, schizophrenics and residents of mental hospitals. He demonstrated that some characteristics of metabolic pattern could be associated with each of these groups.

Griffiths and Wang [11] reported that metabolomics origins are found in the 60's and 70's in the work of the Horning. Horning and Horning published several papers about metabolic profiles determination in urine by Gas Chromatography hyphenated among others to mass spectrometry [12;13]. At the same time, Robinson and Pauling performed a quantitative analysis of urine vapour and breath by gas chromatography [14].

Metabolomics belongs to the “omics” techniques together with genomics, transcriptomics and proteomics that are related to the genome (DNA), the transcriptome (RNA), and proteome (proteins), respectively (Figure 1). The term metabolome (and

obviously metabolomics) was coined on the basis of genome and transcriptome. It appeared for the first time in a publication by Oliver in 1998 [15]. The metabolome reflects past events that include whole metabolism and the interaction with the environment, whereas the genome reflects the real and potential functional information of organism.

Metabolomics/Metabonomics is the analysis of metabolome in a given condition. Both terms can be interchanged. Initially, metabolomics refers to the measurement of the pool of cell metabolites [16] whereas metabonomics describes "the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification" [17;18]. Nicholson's definition underlines the role of two major scientific disciplines used in metabonomics: analytical chemistry and biostatistics. By consistency, we use the term metabolomics in this manuscript. Metabolomics is therefore a data-driven approach, *i.e.* a technology based on the interpretation of information-rich data aimed at complementing the understanding of biological processes [19].

Each individual (from every living species) owns his steady-state equilibrium called homeostasis. Interactions with the environment (exposure to drugs or chemicals) or the onset of a disease disrupts this homeostasis at different levels of the biological organization, including the metabolome. The concentrations of endogenous metabolites may be altered and xenometabolites may appear. Whereas the latter are obviously markers of exposure (biomarker of exposure for instance), specific signatures of disease or exposure (often referred to as metabolomic profile) could be found by the subtle analysis of endogenous metabolites.

Biological markers or biomarkers are measurable internal indicators of molecular and/or cellular alterations that may appear in an organism after or during exposure to a toxicant and possible disease [20;21]. This definition is used in environmental and occupational toxicology and is larger than that of the National Institute of Health (NIH) that focuses on drug development and defines a biomarker as "a characteristic that is objectively

measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacological processes to a therapeutic intervention” [22].

Biomarkers could be divided into several categories that include biomarkers of exposure, biomarkers of effect and biomarker of susceptibility. Biomarkers are compounds or a set of compounds (metabolomic profile) that must be quantitatively, sensitively, specifically, and easily measurable on non-invasively collected biological media [23]. A biomarker of exposure is an indication of the occurrence and extent of exposure. It depends on the chemical fate of the exposed toxicant in the body. The biomonitoring of exposure has been used for a long time in occupational settings *e.g.* for the determination of lead [24] or benzene metabolites [25] in blood or urine.

Biomarkers of (biochemical) effect(s) indicate that exposure has resulted in an interaction between the toxicant and a biological target. Mutagenic and carcinogenic substances that possess electrophilic function(s) bind to macromolecules such as proteins, DNA or lipids. Hemoglobin is often used in biomonitoring because of its long life span and ease of access. Oxidative stress perturbs the homeostasis of cell and leads to the production of specific substances such as 8-Hydroxy-2'-deoxyguanosine or to an imbalance of glutathione pathway [26].

Biomarkers of susceptibility describe inter-individual differences in response to toxicants from genetic causes or from non genetics factors (age, liver disease, kidney disease, diet, dietary supplementation...). Polymorphisms of activating/detoxifying enzymes have been identified as key factors in the relationship between external (*e.g.* ambient air) and internal exposure (*e.g.* urinary excretion). Haufroid et al [27] demonstrated the relationship between the urinary excretion of phenylhydroxyethylmercapturic acids (a mercapturic acid metabolite of styrene) and the genetic polymorphism of glutathione S-transferase M1. A similar approach, referred to as pharmacometabolomics, has already been proposed to study

the response to drugs [28;29]. In this particular context, metabolomics acts as a functional genomics tool.

How to measure metabolites?

Because metabolites exhibit a high chemical diversity, ranging from sugars to lipids, it is impossible to perform their analysis in biological media with a single and universal technique. The two main analytical platforms which provide structural information relevant for metabolite identification rely on nuclear magnetic resonance (NMR) [30], or on mass spectrometry with different ion sources and mass analyzers [31-40]. Each of these tools provides complementary but sometimes redundant information, as emphasized by Lindon and Nicholson [41]. Beside NMR and Gas chromatography which were pioneering techniques for metabolomics, liquid chromatography hyphenated to mass spectrometry (LC/MS) has emerged as a popular and powerful tool, as shown in figure 2.

Nuclear Magnetic Resonance (NMR) was one of the first method used for metabolomics [41-44]. It is a non destructive, rapid, and highly robust technique which produces highly informative structural information. However, NMR is less sensitive than mass spectrometry and requires, therefore, larger amounts of samples. NMR is often used without any prior separative method and does not require development as is the case with chromatography. However, as each metabolite participates to the NMR spectra, the deconvolution of signals is often a tedious process.

The development of LC/MS significantly impacted biological research, including metabolomics. Initially, gas chromatography was the only separative method able to be hyphenated to mass spectrometry. However the use of gas chromatography is restricted to a small set of biological molecules, *i.e.*, those that are volatile or could be derivatized. As a consequence, biological molecules of high molecular weight, such as proteins or nucleic acids, were excluded. The situation was improved by the introduction of atmospheric pressure

ionization mass spectrometry (API-MS)-based techniques combined with liquid chromatography which exhibit a good sensitivity, high dynamic range and versatility but also provide soft ionization conditions giving access to the molecular mass of intact biological molecules.

One of the strengths of API-MS-derived tool is the high diversity of analyzers available: triple quadrupoles, ion traps, time of flight, Orbitrap and Fourier transform-ion cyclotron resonance instruments, the three latter providing high resolution and accurate mass measurements. Among these technologies, high resolution analyzers are becoming increasingly popular in the field of metabolomics because they provide (i) accurate mass measurement, which are useful for the determination of elemental composition of metabolites, and (ii) structural information with MS/MS or sequential MSⁿ experiment, especially when ion products are analyzed at high resolution.

The aim of this paper is to review the metabolomic approach for biomarker discovery in the field of toxicology and clinical chemistry, by focusing on the use of LC/MS. We will successively describe the overall process of analysis (*i.e.*, data acquisition, statistical analyses and metabolites identification), review some seminal papers in the field and discuss the perspectives of metabolomics for biomonitoring of exposure and diagnosis of diseases.

LC/MS based metabolomics: a practical approach

A metabolomics experiment starts with an appropriate experimental design ensuring that the data will be relevant for further biological interpretation. The experimental step begins with the treatment of the biological samples before injection into LC/MS systems. The resulting metabolic fingerprints are then pre-processed using automatic peak detection softwares before being analyzed with appropriate statistical tools. Finally, identification of the discriminating signals is undertaken by combining mass spectrum analysis, database

consultation and other mathematics and informatics tools. All these critical steps are displayed in the Figure 3 and will be detailed in this section.

Design of the experiment

One of the issues in metabolomics is the occurrence of confounding factors that mask the biological phenomenon to be investigated. These confounding factors can be of either analytical or biological origin and their impact has to be anticipated as far as possible by properly designing the experiment.

Many putative confounding factors of biological origins have already been pointed out in published papers: age [45], gender [46], chronobiological effects [47;48], animal species and strains [49;50] and even environmental factors such as diet and gut microflora [51-53]. Some of these factors such as diet vary from one subject to another and cannot be easily controlled; the only possible option is to keep their presence in mind. Other factors such as age and gender should be balanced throughout the different groups to limit their impact on further statistical analyses. In this context, it is of special interest to investigate the metabolic profiles recorded from biofluids of “normal” healthy subjects in order to evaluate the impact of these physiological factors on the metabolite levels [54;55]. This should underpin the use of LC/MS based metabolomics in the clinical chemistry and toxicology arenas.

Another important issue is the normalization of results. This is especially the case for urine samples. Indeed, contrary to most biological fluid or tissue samples, in which metabolites concentration is clearly related to volume or quantity drawn, urinary metabolites concentrations are very fluctuating because of urine volume and clearance variations. Thus, normalization of data is necessary to compare urinary metabolic profiles. It can be performed by weighting the signal abundances in each sample by the urinary volume, creatinine concentration, osmolality or total useful MS signal recorded from mass spectra [56],

according to the type of sample (spot urine or 24h collection) and information available (urinary creatinine concentration or volume, for example).

Beside these biological confounding factors, analytical issues also have to be considered. This is for example the case with the clogging of the electrospray source which progressively alters the detection of analytes. This leads to a clear discrimination between samples analyzed at the beginning and at the end of an experiment which could hamper the visualization of the biological effect of interest. A way to address this issue is to randomize the samples throughout the sequence of injections.

Sample preparation

Metabolite extraction strongly depends on the type of biological medium (*i.e.*, cell extracts or biofluids), and also on the chemical structures of the metabolites to be preferably detected (*i.e.*, polar compounds or lipids). Urines samples are often just diluted with water before injection into the LC/MS system [57-59], whereas other protein-rich biofluids such as plasma or cerebrospinal fluids are processed using organic solvents such as methanol, ethanol or acetonitrile [60;61]. The same kind of procedures may be applied to cell samples: after having been centrifuged to separate cells from supernatant, the cell pellet may be resuspended in water/cold organic solvent mixtures and then sonicated or mechanically agitated to disrupt cell membranes [62]. Tissues have first to be quickly collected and frozen by plunging them in liquid nitrogen for example. This is followed by homogenization in cold organic solvents such as methanol. A Folch derived extraction protocol can then be used either to clean the polar fraction from insoluble lipids or to analyze both polar and apolar fractions [63;64]. Then, depending on the type of organic solvent used for metabolite extraction, samples are diluted in the mobile phase or centrifuged, evaporated to dryness and finally resuspended in a solvent compatible with further injection into the LC/MS system.

Acquisition of metabolic fingerprints using LC/ESI-MS systems

Initially, the acquisition of metabolic fingerprints was performed using LC coupled to electrospray mass spectrometers equipped with low resolution detectors such as triple quadrupole [65;66] or ion trap [67] analyzers. By these means, it was possible to separate and detect thousands of ions in biofluid samples. However, the interpretation of data was limited by both insufficient chromatographic separations and also identification issues. Indeed, it was difficult to link an experimental mass measured at low resolution and low accuracy to a metabolite among many others having the same nominal mass. These two issues have been partly addressed by the implementation of (i) ultra performance liquid chromatography (UPLC), which improved chromatographic resolution, peak capacity, and even sensitivity [39] and (ii) high-resolution mass spectrometers such as time-of-flight (TOF) and Fourier transform (FT) mass spectrometers.

High- and ultra-high-resolution analyzers are becoming increasingly popular in the field of metabolite profiling because they provide accurate mass measurements which are useful for the discrimination between isobaric ions, and even isomers if their fragmentation patterns are different [68], leading to the detection of a higher number of signals than that obtained with low-resolution analyzers. Of course, accurate mass measurements also enable the determination of elemental compositions of metabolites for further identification.

Finally, it is important to check for the consistency of analytical results before biological interpretation. Indeed ion abundances can decrease for long-term analysis (intra-experiment variability), but also from an experiment to another (inter-experiments variability) because of the degradation of MS or chromatographic separation performances [69]. This complicates the automatic detection and alignment of features, and also stitching together of datasets. A normalization step is thus required. To this end, a mixture of reference compounds can be injected at regular intervals to assess the performances of both the

chromatographic column (i.e., consistency of retention times and peak widths of reference compounds) and the mass spectrometer during the experiments (i.e., consistency of mass accuracy and signal intensity of reference compounds). The same kind of approach can also be performed by using quality control (QC) samples that are representative of the biological samples to analyze [70], but this is not sufficient to normalize peak intensities. That is why many normalization approaches have been developed to overcome analytical variability, such as NOMIS (Normalization using Optimal selection of Multiple Internal Standards) or CCMN (Cross-Contribution Compensating Multiple Standard Normalization) [71], and also to facilitate comparison of datasets [72;73].

Automatic detection of ions

The aim is to represent the initial raw data in a matrix format which is compatible with subsequent statistical and biochemical analyses. As data formats are proprietary, a conversion step into universal data formats such as netCDF (Network Common Data Form - www.unidata.ucar.edu/software/netcdf/) [74] or mzXML [75] are required before running the data processing. A typical processing pipeline includes filtering, feature detection, alignment and normalization. This can be achieved by using dedicated commercial, or free and/or open-access software, as already reviewed by Katajamaa et al. [76]. In the latter case, it is possible to have access to the algorithm and to modify or improve them. This is for example the case for XCMS [77;78] and MZmine [79].

These software tools also differ by the implemented approaches. While the subtraction of the background noise often relies on filtering algorithms classically used in signal processing, large differences are observed at the level of the detection and alignment of the signals. As an example, the detection of peaks is achieved in both the retention time and m/z dimensions in an independent way with the software *Mzmine* (<http://mzmine.sourceforge.net/>) [79], whereas the *MatchedFilter* algorithm of the XCMS software

detects the ions from m/z windows [77]. As an alternative of binning approaches, the *centWave* algorithm, which is also part of the XCMS software, performs a two-dimensional feature detection by using a combination of a density based technique to detect regions of interest in the m/z domain, and a Wavelet based approach to resolve chromatographic peaks [78]. At the opposite, other tools such as *MathDamp* ([http:// mathdamp.iab.keio.ac.jp/](http://mathdamp.iab.keio.ac.jp/)) [80] make a comparative analysis from the original data without any signal detection step.

Limited information about the validation of signal detection softwares is available in the literature. The main reasons are that most of the signals present in the metabolic fingerprints remain uncharacterized and the results of an automatic detection procedure of metabolites may be impacted by both the type of instrument and the biological medium. As a consequence, users have to evaluate the software with their own criteria in order to select that or the most suited one(s) regarding their instruments and the biological matrix.

Tautenhahn et al. proposed an interesting approach to evaluate signal extraction software. It is based on the estimation of three parameters: the recall, which measures the fraction of relevant features that are extracted by the algorithm, the precision, which is the percentage of relevant items compared with the false positives, and the run time, which is the time required for the algorithm to achieve feature detection from a given data set [78]. Actually, many artefactual signals are present in the data matrices following automatic data extraction and signal alignment. They have been evaluated as around 400 for 100 relevant features [78]. One way to address this issue is to perform serial dilution of QC samples and select the features whose levels are correlated to the dilution factor [62;78].

Statistical analyses

As for transcriptomics or proteomics, metabolomics relies on differential analyses of metabolic fingerprints which lead to a semi-quantitative expression of the results (*i.e.*, decreased or increased area or intensity ratios). As it appears difficult to handle and to

compare data sets which contain several hundreds to thousands of signals, multivariate statistical analyses are required to address this issue [81] (Figure 4).

Data exported from automatic peak detection software tools have first of all to be scaled. A typical procedure relies on unit variance scaling: the variables are centered and divided by their standard deviation. This gives an equal weight to signals exhibiting very different abundances. However, in some case, this may lead to a dilution of the analytical information of biological relevance and other methods such as pareto-scaling (the variables are centered and divided by the square root of their standard deviation) may be preferred.

Once the data have been scaled, a preliminary step often relies on the use of unsupervised analyses such as principal component analysis. This descriptive method does not require any information about the nature of samples. It enables to visualize the organization of the original data in a two or three dimensional space by reducing the dimensionality of complex data sets. Explicative analyses are then performed by using supervised tools such as PLS (projection to latent structures or partial least squares) regression, PLS-discriminant analysis (PLS-DA), or more recently OPLS (Orthogonal Projection on Latent Structure) in order to facilitate the isolation of the ions responsible for the discrimination between groups [82].

Finally, a clear distinction has to be done between exploratory studies that try to reveal new biomarkers whose biological relevance has to be established, and predictive studies that aim at classifying unknown subjects and for which the issues of statistical powerfulness and validation are critical [83].

Identification

The metabolite identification process using atmospheric pressure ionization mass spectrometry-based tools starts with the interpretation of the mass spectra in order to ensure that the signal of interest really corresponds to a monoisotopic ion and not to an isotope,

adduct or ion product generated during the ionization process. Several informatics and mathematics tools are available for that purpose. They are grouping the signals related to given metabolites according to (i) specific mass differences corresponding to isotopes, adducts, and product ions, and (ii) the correlations between the intensities of pairs of ions, either across several spectra within a sample or across all samples where the signals are observed [59;84;85] (Figure 5).

One or few relevant elemental composition(s) is/are deduced from accurate mass measurements if high or very high resolution mass spectrometry is available, for further database queries. Collision induced dissociation (CID) spectra are then acquired and interpreted in order to get information about the chemical structure. At this stage, chemical database queries may be refined and the highlighted compounds, if any, are kept for further consideration or ruled out based on chromatographic retention time and CID mass spectra information. Complementary experiments (*i.e.*, other sequential MSⁿ experiments or H/D exchanges) may be required before obtaining or synthesizing the reference compounds. Finally, formal identification is achieved when the metabolite to be characterized exhibits the same retention time and CID spectra than those of the reference molecule.

In the Metabolomics Standards Initiative [86], Sumner and al. have reported four different levels of identification according to the information provided :

(i) Identified compounds: a minimum of two independent and orthogonal types of data relative to an authentic compound analyzed under identical experimental conditions. In MS-based techniques this could include: retention time/index and mass spectrum, or accurate mass and tandem MS.

(ii) Putatively annotated compounds: without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries.

(iii) Putatively characterized compound classes: based upon characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class.

(iv) Unknown compounds: although unidentified or unclassified these metabolites can still be differentiated based upon spectral data, thus enabling relative quantification.

Mass spectrometry experiments alone may be sufficient when the metabolites to characterize are well described in databases, commercially available and discriminated from isomers thanks to an adequate chromatographic separation and/or characteristic MSⁿ spectra. In this case, identification is achieved by matching the retention time and CID spectra of the compound of interest to those of the putatively related synthetic reference molecule. However, in many cases, the metabolites of interest are not reported in any biochemical or metabolomic databases and additional analytical tools such as NMR cannot be used due to a lack of sensitivity and/or insufficient chromatographic separation. The only solution is then to perform a careful and precise interpretation of CID spectra combined with additional experiments such as H/D exchange in order to provide new structural hypotheses that have to be assessed by further chemical synthesis [85].

Selected applications in the field of toxicology

Toxicology aims at studying adverse effects of chemicals (xenobiotics) on living organisms. The toxicity of a given compound refers to its ability to disrupt some biological functions at a certain level of biological organization (*i.e.*, cell, tissue, or organ). It is related to the amplitude and the duration of the exposure and also to the degree of absorption of the substance by the organism, its distribution, biotransformation and elimination or accumulation. Understanding the mechanism of a toxic event is a challenging task, especially in the field of drug research and development. Indeed, target organ toxicity remains an issue and idiosyncratic toxicity, which refers to individual susceptibility in drug induced toxicity, is

often not detected before the drug has been on the market (Rofecoxib [87], Rimonabant [88]). Many *in vitro*, cell and animal models are designed to address these issues, but they may not be easily extrapolated to human. Biomarkers are useful to predict a toxic event before the occurrence of clinical events (biomarkers of early effect), to evaluate the severity of the poisoning (biomarkers of effect), and also to monitor exposed patients (biomarkers of exposure). This is another challenge because the occurrence of adverse effects has multiple origins including host environment interactions that are difficult to be caught using conventional approaches for biomarker discovery which are focused on limited biochemical and metabolic aspects.

By achieving a global detection of molecular events at the different levels of biological organization, omics approaches may provide answers to these issues, as emphasized by early proof-of-concept studies in toxicogenomics [27], transcriptomics [89] and proteomics [90]. Metabolomics, which enables to track homeostatic disruptions and host-environment interactions, is of particular interest in this context. Pioneering studies using NMR have already been published and also reviewed [17;19;44;91-96], and the consortium on metabonomic toxicology (COMET), coordinated by the Imperial College and including pharmaceutical companies, has started to develop expert models for the classification of toxicity based on ¹H-NMR analysis [19]. However, none of them have ever been published until now. The development of LC/MS in this field is relatively recent. Several publications illustrating metabolomics applications in the field of toxicology are displayed in the table 1. They address biomarker discovery, predictive models and mechanistic considerations mainly in the field of hepato- and nephrotoxicity by using model toxicants. The input of LC/MS based approaches will be reviewed and discussed in this section.

LC/MS based metabolomics for toxicity biomarkers discovery.

Many studies are performed using different analytical platforms, such as ¹H-NMR, GC/MS and LC/MS in order to maximize the metabolite detection coverage. Most studies attempt to address the issue of organ toxicity and aim at finding metabolite concentration changes related to the toxicant, occurring before clinical or histopathological detections and being more specific than conventional biomarkers such as alanine aminotransferase (ALT) and aspartate aminotransferase (AST) enzyme activities or bilirubin for hepatotoxicity, or blood urea nitrogen (BUN) for nephrotoxicity.

Among these studies, acetaminophen (also known as *N*-acetyl-*p*-aminophenol, APAP) is frequently used as a model drug for hepatotoxicity. It is cleared from the body through hepatic glucuronide and sulphate conjugation. However, in case of overdose, these metabolic pathways are saturated and reactive metabolites such as *N*-acetyl-*p*-benzoquinone imine (NAPQI) are produced. NAPQI reacts with glutathione (GSH) to form a conjugate, which is subsequently degraded to a mercapturic acid derivative that can be detected in urine. However, NAPQI can also oxidize glutathione and in turn be reduced back to paracetamol [97]. When the GSH pool is depleted, NAPQI reacts with cell macromolecules. This mechanism is supposed to be one of the explanations for hepatic necrosis recorded in cases of APAP poisoning.

Sun et al. (2008) investigated the acute and chronic toxicity of acetaminophen on male Sprague-Dawley rats by metabolomics using NMR and UPLC coupled to an electrospray Q-TOF mass spectrometer [98]. Metabolic changes were matched up with histopathological observations and others markers of liver injury (serum ALT, AST and bilirubin) to highlight metabolites related to APAP induced toxicity. Necrosis was not observed in the course of the chronic study and was only detected at the highest dose (*i.e.*, 1600 mg/kg) of the acute study at the 48 h time point. Urinary metabolite concentration changes were observed in both acute

and chronic studies from the 400 mg/kg dose. Both NMR and UPLC/MS pointed out depletions of antioxidants and energy metabolites. The decrease of 1-methylnicotinate levels observed by NMR was of particular interest because this molecule is linked to the glutathione biosynthesis pathway (this metabolite is produced during the conversion of S-adenosylmethionine to S-adenosylhomocysteine). The decrease of urine 1-methylnicotinate concentration could be related to the depletion of S-adenosylmethionine observed by the authors using a targeted LC/MS/MS assay.

These results have been confirmed by another study that focused on the regulation of the trans-sulfuration pathway in liver toxicity conditions using NMR, LC/MS, and also gene expression data: the expression of genes involved in the trans-sulfuration pathway was decreased in the liver, whereas taurine, creatine (observed by NMR) and S-adenosylmethionine (observed with LC/MS) levels were increased in urine following APAP administration to rats [99].

One of the strength of LC/MS-based metabolomics is the possibility to detect many xenobiotic related metabolites thanks to its high sensitivity. Sun et al. detected 6 APAP metabolites in rat urine and concluded that approximately 95 and 65 ions were related to APAP metabolites in negative and positive modes, respectively. They decided to remove these signals in order to facilitate the observation of endogenous metabolites whose levels were altered following APAP administration [98]. However, some of these so-called xenometabolites can also provide the toxicologist with mechanistic information about drug toxicity and thus being used as biomarkers. In a following study, Sun et al. [100] investigated the excretion kinetics of APAP metabolites in rat urine and observed that the concentrations of the APAP-N-acetylcysteine conjugate exhibited a significant correlation with AST activity, bilirubin, creatine and histopathological observations, and a significant anticorrelation with S-

adenosylmethionine levels, suggesting that it is a good indicator of APAP-induced liver injury.

Although urine is the biofluid of choice for metabolomics (easily sampled, simple to analyze and providing investigators with information about polar metabolites including energy metabolites and xenometabolites) complementary information about lipids can be obtained from other biological media such as plasma. For example, using a LC/MS-based metabolomic approach, Chen et al. detected an accumulation of long chain acylcarnitines in plasma from APAP treated mice [101]. This was reinforced by the concomitant observation of increased free fatty acids and triglycerides plasma levels using colorimetric assays. Thanks to additional experiments performed on CYP 2E1 and PPAR α null mice, the authors concluded that inhibition of fatty acid β -oxidation through the suppression of PPAR α activation is a contributing mechanism of APAP-induced hepatotoxicity and that long chain acylcarnitines could be early biomarkers of APAP hepatotoxicity that may complement the measurements of GSH levels and serum AST or ALT activities.

LC/MS-based metabolomics for the building of predictive models of toxicity.

Beside biomarker discovery, other studies report the development of metabolomic-based approaches for predicting and classifying different modes of toxicity. This is for example the case with La et al. in the field of chemical-induced hepatotoxicity [102]. They applied LC/MS to analyse urine samples of rats treated with four different hepatotoxins: R-naphthyl isothiocyanate (ANIT), carbon tetrachloride (CCl₄), APAP, and diclofenac. They found specific patterns of metabolites concentration changes that were characteristic of each hepatotoxin and managed to build a mathematical model exhibiting predictability higher than 95% by using linear discriminant analysis and soft independent modelling of class analogy with residual distance. However, it is challenging to determine whether these patterns of metabolite concentrations are specific of a mode of organ toxicity (*i.e.*, necrosis or

cholestasis), or rather of compounds or chemical families. To address this issue, several toxicants exhibiting different chemical structures, but the same mode of organ toxicity should have been included in the experimental protocol. Such an approach has been performed by Boudonk et al. in the field of nephrotoxicity.

Boudonck et al. [103] reported on a metabolomic investigation on 3 drugs (gentamicin, cisplatin and tobramycin) inducing proximal tubule nephrotoxicity. Urine and kidney were collected after one, five and twenty-eight dosing days and the samples were analyzed using GC and LC/MS. About 30% and 70% of the metabolites observed in kidney extracts were detected by GC/MS and LC/MS, respectively, whereas half of them were measured by both techniques in urine. Increases in amino-acids and polyamines were observed in urine and decreases in purine and pyrimidine nucleosides were detected in kidney tissues before observable kidney injury by conventional histology and clinical chemistry tools. Urinary metabolites exhibiting significant changes with all 3 drugs, such as branched chain amino-acids, hippurate and glucose at day 28 were then selected to build a predictive model based on classification trees in order to predict the onset of the nephrotoxicity at days 1 and 5.

Van Vliet et al. developed an *in vitro* model to evaluate neurotoxicity based on rat primary re-aggregating brain cell cultures followed by LC/MS-based metabolomics [104]. Cell cultures were exposed to the neurotoxic methyl mercury chloride (MMC) at concentrations ranging from 0.1 to 100 μ M or to the brain stimulant caffeine at concentrations ranging from 1 to 100 μ M. The occurrence of cytotoxicity was assessed by the detection of an increased activity of the lactate dehydrogenase (LDH) in the culture media. No neurotoxicity was observed with caffeine, whereas it occurred from 1 μ M with MMC. Interestingly, differences in metabolite concentrations were observed between control and MMC exposed sample, as emphasized by concentration dependent clusters observed on principal component

analysis score plots. The concentration of five metabolites was either increased (creatine, spermine) or decreased (glutamine, GABA and choline) in MMC samples, and these metabolites were found responsible for this clustering. To evaluate their model, van Vliet et al. tested 8 compounds exhibiting different modes of organ toxicity. On the PCA score plots, controls and hepato- and nephrotoxic compounds were part of the same cluster whereas neurotoxic compounds were clearly individualized.

Toward mechanistic considerations

Williams et al (2005) showed the value of using MS-based metabonomics for elucidating the mechanism of toxicity of D-serine in rat. D-serine is a nephrotoxic amino acid that causes selective necrosis of renal proximal tubule cells in rats [105] by an unknown mechanism. Using ¹H-NMR, Williams et al. [106] showed that D-serine-induced kidney tubular damage was associated with proteinuria, glucosuria and aminoaciduria, which have already been described as unspecific markers of tubular nephrotoxicity. Further LC/MS analyses led to the identification of several metabolites (hydroxypyruvate, glycerate, sebacic, xanthurenic and methyl succinic acids, acyl carnitine) that had not previously been detected by ¹H NMR [107]. Interestingly, glycerate and hydroxypyruvate are produced from serine by the peroxisomal enzyme D-amino-acid oxidase (D-AAO). The authors hypothesized that hydroxypyruvate generates hydrogen peroxide, which induces a peroxisomal oxidative stress. The resulting peroxisomal dysfunction leads to decreased fatty acid metabolism and oxidation, as emphasized by the observation of decreased levels of the dicarboxylic acids such as sebacic and methylsuccinic acid, and acylcarnitins. Of note, a relationship between tryptophan catabolism and peroxisomal metabolism has already been reported [108] and is consistent with the perturbations observed in this study: increased excretion of tryptophan and decrease of xanthurenic acid and other TCA cycle intermediates. Finally, it has been shown in another study that the co-administration of D-serine and sodium benzoate, a potent

competitive inhibitor of renal D-amino-acid oxidase, prevents kidney injury, thus confirming the implication of this enzyme in the mechanism of toxicity [109].

Finally, new insights into mechanisms of toxicity will probably be obtained through the coupling of classical histological and biochemical tools with other more recent omics and imaging approaches. In this context, a collaborative research effort in molecular system toxicology has been launched by the FDA's National Center for Toxicological Research and BG Medicine Inc. It is supported by 7 pharmaceutical companies and 3 technology providers and aims at investigating drug induced liver toxicity. Three days and twenty-eight days dosing studies are performed on related compound pairs, including a "clean" compound and a toxic one in order to highlight off-target molecular responses. Proteomics, LC/MS-based metabolomics and gene expression data are obtained from liver extracts, proteomics and LC/MS based metabolomics are obtained from plasma samples, and NMR based metabolomics experiments are performed on urine samples. These data are then confronted with histology and classical clinical chemistry tools. Preliminary findings are reported in a publication in which the performances of the analytical platforms on a first compound pair (entacapone and tolcapone) are presented and discussed [110].

Selected applications in the field of clinical chemistry

Clinical chemistry deals with any analysis performed on body fluids for medical purpose, including disease diagnosis and follow-up, and also therapeutic drug monitoring. MS-based approaches are used in clinical laboratories since the 1970s [111]. Most of them are targeted methods focusing on particular metabolites or chemical families. Currently existing tandem MS methods are used to carry out neonatal screening analysis using the same principles as metabolomics {American College of Medical Genetics/American Society of Human Genetics Test and Technology Transfer Committee Working Group, 2000 459 /id}. This is ultimately the tangible use of MS-derived discovery of novel biomarkers. Now,

technological and bioinformatical improvements of this last decade have enabled the implementation of MS-based global approaches, namely metabolomics, in the field of clinical chemistry. Table 2 displays some key applications of MS-based metabolomics to clinical chemistry. They address different medical areas such as cardiology, transplantation, human reproduction, diabetes, central nervous system diseases, or oncology.

Thanks to the versatility of API-MS-based tools, MS-based metabolomics offers the possibility to provide chemists and physicians with various snapshots of different biological media such as plasma, urine, cerebrospinal fluids, cell extracts or tissue extracts obtained from biopsies, as shown in table 2. These snapshots may focus on concentration changes of selected metabolites sometime occurring at trace levels. Triple quadrupole mass spectrometers operated in the selected reaction monitoring mode are the instruments of choice for such targeted approaches thanks to their sensitivity. Snapshots may also focus on particular metabolite families such as carnitine species or lipids. In the first case, the selectivity of the detection may be brought by MS/MS detection, thanks to the constant neutral loss or parent ion scanning modes that are available on triple quadrupole instruments, whereas, in the second situation, it is rather obtained by specific sample treatment procedures, such as Folch extraction [112]. At last, the pictures may provide the user with an overview containing many features related to both genetic and environmental contributions (*i.e.*, diet, lifestyle, gut microbial activity, drug intake, and exposure to pesticides, plasticizers or food preservatives...). The terminology of global approach has been coined for this kind of picture and high and ultra-high resolution mass spectrometers are the most frequently used instruments in this context. These aspects will be discussed in this section with selected applications.

In targeted approaches, metabolites are selected with regards to their biological relevance to the field of investigation, or because they are representative for known metabolic

pathways. Metabolites exhibit very different structures and several analytical methods may be used in parallel for their detection. Sabatine et al. [113] and also Lewis et al. [114] used 3 LC columns for sugars and nucleotide, organic acids and amino-acid analyses, whereas Turer et al., used colorimetric methods to detect glucose, lactate, free fatty acids and pyruvate, and a MS-based method to profile amino-acids and carnitine derivatives [115]. Several tens to hundreds of metabolites may be detected using MS/MS analysis performed on triple quadrupole instruments in most cases. These metabolites are mainly amino-acids, sugars, organic acids involved in central and energetic metabolism, and also lipids (carnitine derivatives and phospholipids).

Shaham et al. conducted a 3 steps study to highlight metabolic disorders linked to respiratory chain dysfunction [116]. They started to work on a cell model of chemical inhibition of the respiratory chain by applying LC-MS/MS using a triple quadrupole device and also ¹H NMR. Among the 191 detected metabolites by LC-MS/MS, the levels of 32 of them were found to be altered by the RC inhibition. This was for example the case for alanine, lactate (both current biomarkers of RCD), glucose, creatine and others TCA-cycle intermediates. In the second step, they screened human plasma from two cohorts of patients with pathogenic mutation or abnormally low respiratory chain enzyme activity in muscle using the LC-MS/MS method. Concentration trends similar to those observed in culture media were found in the first cohort (16 patients and 25 controls) for 26 of the 32 metabolites previously pointed out. The levels of lactate (+107%), alanine (+46%), creatine (+233%) and uridine (-24%) were significantly different between patients and controls. To confirm the huge increase in creatine in plasma samples from patients suffering of a respiratory chain dysfunction, a second independent cohort (14 patients and 4 controls) was analyzed using another LC-MS/MS method. Increased concentrations of creatine (201%) were observed in patients whereas alanine and lactate levels did not significantly differ. This study suggests that

plasma creatine levels are more powerful biomarkers of respiratory chain dysfunction than alanine or lactate. Actually, creatine (Cr) is converted to phosphocreatine (PCr) by creatine kinase in the presence of ATP and the PCr/Cr balance reflects the cell energy state. In case of respiratory chain dysfunction, the low energy state makes that low intracellular ATP and PCr concentrations are observed, together with an accumulation of intracellular Cr which leads to an increased excretion of Cr in the extracellular medium.

Oresic et al. reported on an ongoing cohort of children who progressed to type 1 diabetes [117]. They analyzed serum metabolites before, around the time and after conversion using an UPLC/Q-TOF for lipid profiling and two dimensional gas chromatography coupled to a TOF mass spectrometer for polar metabolites. Although the mass detection was performed on a large mass range, these authors decided to keep only identified metabolites, *i.e.*, 53 lipids and 75 metabolites, for further statistical analyses and biological interpretation. Results show that children who developed diabetes had metabolic perturbations before and during conversion such as increased levels of lysophosphatidylcholine, GABA, glutamic acid and leucine and decreased levels of succinic acid, PC, ketoleucine and glutamine. Normalization of profiles after the conversion indicates that those metabolic dysregulations precede B cell autoimmunity and disease onset. The interdependence of metabolic and immune system factors raises many questions, including the possible role of choline and intestinal microbiota in lipid dysregulation, or plasmalogen and oxidative damages toward B cells. However, tissue-specific mechanisms behind metabolic disturbances remain unclear: is the autoimmunity a physiological response aimed at restoring the metabolic homeostasis, or do metabolic dysregulations reflect an early stage of this immune response toward the B cell autoantigens that are not yet detectable?

Few papers using untargeted LC/MS based approaches in the field of clinical chemistry have been published. Two of them address the issue of cancer biomarker discovery.

Sreekumar et al. applied MS-based metabolomics in an attempt to identify biomarkers for non-invasive diagnosis and prognostic cancer invasion and disease aggressiveness [118]. To this end, they analyzed with two complementary GC and LC/MS methods the concentrations of several hundreds of metabolites across 262 prostate-related samples (42 tissues extracts, 110 plasma and urine samples from biopsy-positive and negative prostate cancer patients). As amino acid metabolism and methylation of biomolecules were known to be involved in prostate cancer progression, they focused on metabolites of these pathways the levels of which were found increased in tissue samples from cancer patients. This was the case for sarcosine, a N-methyl derivative of glycine, which was significantly increased in tissues during disease progression from benign disease (undetectable) to metastatic stage (higher levels than in organ-confined disease). However, monitoring sarcosine levels in prostate tissues appears of limited interest because histology is already available and more powerful for the diagnosis and the prognosis of cancer. The authors then decided to turn to urine and found out that sarcosine was detectable but occurred at trace levels, with a modest but significant predictive value (more sensitive than PSA) for prostate-cancer diagnosis and disease progression. Beside the discussion about methylation and sarcosine, many other discriminating metabolites, identified or not were not taken into account. One reason for that is the lack of appropriate tools to visualize and synthesize omics data sets.

Richie et al. [119] applied both targeted and non targeted metabolomics-based approaches in order to identify non invasive biomarkers of colorectal cancer (CRC). They started with a large-scale study on serum samples from several populations of different origins. Discriminating metabolites common to CRC patients whatever their ethnic or geographic origin were highlighted using a non targeted FT-ICR MS-based method combined with statistical analyses. A group of metabolites the levels of which were dramatically decreased in the 3 populations was pointed out, and thanks to the high resolution and accurate

mass measurement, the authors attributed a mass formula to each signal and noticed that these molecules belong to the same family. Then, they performed MS/MS experiments using a Q-TOF instrument and identified the molecules of interest to be hydroxylated polyunsaturated ultra long-chain fatty acids. This was further confirmed by NMR experiments and MS/MS analysis of reference compounds. Finally a semi-quantitative method was developed on a triple quadrupole instrument operated in the selected reaction monitoring mode for biomarker validation purposes.

Finally, although results from exploratory studies in the field of clinical medicine are promising, many of them should be cautiously interpreted. Indeed, many published studies involve reduced cohorts. This limits the power of statistical analyses and further isolation of discriminating ions. Furthermore, even though such ions are underlined, their formal identification and the demonstration of their biological relevance will be required for further medical applications.

Conclusion and perspectives

The metabolome is characterized by a large diversity of chemical structures requiring diverse analytical platforms to reach its extensive coverage. NMR has been extensively used since the beginning of metabolomics, whereas the use of LC-MS has progressed and is now very popular because it is versatile, sensitive and brings complementary information about biomolecules such as peptides and lipids. The aim of this review was to introduce LC/MS-based metabolomics and to present and discuss key applications focusing on toxicology and disease biomarkers. Whereas the published applications in the field of toxicology still remain proof-of-concept studies, due to the complexity and multifactorial origin of toxicity, the situation seems different in the field of clinical chemistry for which multiplexed targeted approaches provide the clinician with information on few tens to hundreds of metabolites by using MS/MS analysis performed on triple quadrupole mass spectrometers. Furthermore,

recent improvements in mass spectrometry have improved the efficacy of global approaches by facilitating the identification of metabolites of interest thanks to high resolution and accurate mass measurements.

From a technical point of view, metabolomics is the combination of analytical chemistry, statistics and bioinformatics tools that are used separately or together to perform (i) sample preparation, (ii) acquisition of metabolic fingerprints, (iii) automatic detection of ions, (iv) statistical analyses and (v) identification. However, despite recent technological and conceptual improvements, metabolomics appears to be still in its infancy and each step that is mentioned above is a bottleneck in itself. How to accelerate metabolomics studies is therefore a titillating issue.

Three major pitfalls must be highlighted: (i) analytical issues such as the difficulty to compare experiments one to another (developed in the “*Acquisition of metabolic fingerprints using LC/ESI-MS systems*”), (ii) the amount of information generated by metabolomics and (iii) the lack of chemical repositories designed for metabolomics studies (*i.e.*, a central open source of mass and CID spectra acquired with various instruments in different laboratories) that will be helpful in the identification of discriminating signal.

Metabolomic analysis processes generate, especially at the output of automatic ion detection, large matrix of data containing tens of thousands of variables (m/z –retention time). The reduction of data could be performed by taking advantage of signal redundancy, as previously explained. However, the main part of the information remains, until now, unexploited. The only alternative should be the systematic (and automatic) identification of all signals. This is actually one of the major rate-determining steps of metabolomics. The data sets obtained from high and ultra-high resolution mass spectrometry can be processed by informatics tool for automatic query in metabolic and metabolomic public databases with the measured accurate masses. Although such annotations are useful to start with biological data

interpretation, they have to be confirmed by a careful interpretation of the mass spectra, as shown in figure 6.

The bottom of the figure 6 displays a peak list extract generated by an automatic detection software from a LC/MS based metabolomic analysis of human cell extracts. Putative annotations of signals are provided in the fourth column. They have been obtained by matching the experimentally measured masses with those of metabolites contained in public databases such as HMDB [120], KEGG [121] and Metlin [122]. They indicate the putative presence of spermidine, diaminopropane, 3-buten-1-amine, 1-Methylpyrrolinium and cyclopropylamine in cell extracts. Another annotation using a home-made spectral database confirmed the presence of spermidine, but shows that 3 of the other annotations are erratic because the related ions actually correspond to ion products of spermidine generated in the electrospray source during the desolvation process. This example highlights the complexity of API-MS-based data sets and a thorough inspection of mass spectra requiring spectral libraries is necessary before biological interpretation.

Two different kinds of libraries are available for API-MS-based metabolomics: mass spectral and CID mass spectral libraries. The building of the first ones relies on careful interpretations of mass and CID spectra of reference compounds. They aim at annotating biological datasets, as shown in figure 6, whereas the latter are useful to confirm peak list annotations and to characterize unknown compounds. These libraries should be shared between users in order to make metabolite identification in various biofluids effective. Unfortunately, API-MS exhibits poor reproducibility and high inter-instrument variability in the generation of fragmentation patterns, thus hampering the constitution of universal databases as done with electron ionization mass spectrometry [123] or with NMR [124].

Despite these limitations, databases containing API mass spectra combined with CID spectra such as HMDB, Metlin, mass bank from metabolome.jp, and lipid maps are beginning

to be released. However, the use of such spectra for comparison and identification must be performed carefully and may lead to erroneous results [85]. This issue of spectral comparison begins to be addressed. For instance, Palit et al. proposed a fragmentation energy index for the normalization of collision energy [125] which is nevertheless restricted to ion trap instruments. Oberacher et al. designed a multicenter study in which 22 test compounds (drug standards) were sent to three different laboratories, where 418 tandem mass spectra were acquired using four different instruments from two manufacturers including Q-TOF, triple quadrupole, Q-Trap and FTICR mass spectrometers [126]. CID mass spectra were recorded without any standardization of experimental conditions and they were matched against a reference library using a sophisticated matching algorithm [127]. The high percentage of correct assignments suggests that it is possible to compare CID spectra obtained from different instruments and laboratories. The possibility of sharing CID spectral libraries and also MS data set repositories should improve the characterization of unknown metabolites of toxicological and clinical relevance.

Acknowledgments

AR is supported by a grant from the Commissariat à l'Energie Atomique (CEA). JFH is supported by a grant provided by the DIANE (Désordres Inflammatoires dans les Affections Neurologiques) project (Région wallonne, Belgium).

References

- [1] Herbert RB. The biosynthesis of secondary metabolites. 2nd ed.: Chapman and Hall, 1989.
- [2] Shargel L, Yu A. Applied Biopharmaceutics and Pharmacokinetics. 4th ed.: McGraw-Hill, 1999.
- [3] Holmes E, Loo RL, Cloarec O, et al. Detection of urinary drug metabolite (xenometabolome) signatures in molecular epidemiology studies via statistical total correlation (NMR) spectroscopy. *Anal Chem* 2007;79: 2629-40.
- [4] Wishart DS. Metabolomics: applications to food science and nutrition research. *Trends in Food Science & Technology* 2008;19: 482-93.
- [5] DeCaprio AP. Biomarkers: Coming of Age for Environmental Health and Risk Assessment. *Environ Sci Technol* 1997;31: 1837-48.
- [6] Wolfe DA. Insights on the utility of biomarkers for environmental impact assessment and monitoring. *Human and Ecological Risk Assessment* 1996;2: 245-50.
- [7] Crockford DJ, Maher AD, Ahmadi KR, et al. 1H NMR and UPLC-MS(E) statistical heterospectroscopy: characterization of drug metabolites (xenometabolome) in epidemiological studies. *Anal Chem* 2008;80: 6835-44.
- [8] Gates SC, Sweeley CC. Quantitative metabolic profiling based on gas chromatography. *Clin Chem* 1978;24: 1663-73.
- [9] Williams RJ. Individual Metabolic Patterns and Human Disease: An Exploratory Study Utilizing Predominantly Paper Chromatographic Methods. Austin: Univ. Texas, 1951.
- [10] Williams RJ. Biochemical Individuality.: John Wiley & Sons, 1956.
- [11] Griffiths WJ, Wang Y. Mass spectrometry: from proteomics to metabolomics and lipidomics. *Chem Soc Rev* 2009;38: 1882-96.
- [12] Horning EC, Horning MG. Metabolic profiles: gas-phase methods for analysis of metabolites. *Clin Chem* 1971;17: 802-9.
- [13] Horning EC, Horning MG. Human metabolic profiles obtained by GC [gas chromatography] and GC/MS [gas chromatography/mass spectrometry]. *J Chromatogr Sci* 1971;9: 129-40.
- [14] Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proc Natl Acad Sci U S A* 1971;68: 2374-6.
- [15] Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. *Trends Biotechnol* 1998;16: 373-8.

- [16] Fiehn O. Metabolomics--the link between genotypes and phenotypes. *Plant Mol Biol* 2002;48: 155-71.
- [17] Beger RD, Sun J, Schnackenberg LK. Metabolomics approaches for discovering biomarkers of drug-induced hepatotoxicity and nephrotoxicity. *Toxicol Appl Pharmacol* 2010;243: 154-66.
- [18] Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 1999;29: 1181-9.
- [19] Lindon JC, Nicholson JK, Holmes E, et al. Contemporary issues in toxicology the role of metabonomics in toxicology and its evaluation by the COMET project. *Toxicol Appl Pharmacol* 2003;187: 137-46.
- [20] Bennett DA, Waters MD. Applying biomarker research. *Environ Health Perspect* 2000;108: 907-10.
- [21] Paustenbach DJ. The practice of exposure assessment. In: Hayes AW, Ed. *Principles and Methods of Toxicology*. London: Taylor and Francis; 2001.
- [22] Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;69: 89-95.
- [23] Timbrell JA. Biomarkers in toxicology. *Toxicology* 1998;129: 1-12.
- [24] Kehoe RA, Thamann F, Cholak J. Lead absorption and excretion in relation to the diagnosis of lead poisoning. *Journal of Industrial Hygiene and Toxicology* 1933;15: 320-40.
- [25] Pearce SJ, Schrenk HH, Yant WP. Microcolorimetric determination of benzene in blood and urine. 1936.
- [26] Angerer J, Ewers U, Wilhelm M. Human biomonitoring: state of the art. *Int J Hyg Environ Health* 2007;210: 201-28.
- [27] Haufroid V, Jakubowski M, Janasik B, et al. Interest of genotyping and phenotyping of drug-metabolizing enzymes for the interpretation of biological monitoring of exposure to styrene. *Pharmacogenetics* 2002;12: 691-702.
- [28] Clayton TA, Lindon JC, Cloarec O, et al. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature* 2006;440: 1073-7.
- [29] Clayton TA, Baker D, Lindon JC, Everett JR, Nicholson JK. Pharmacometabonomic identification of a significant host-microbiome metabolic interaction affecting human drug metabolism. *Proc Natl Acad Sci U S A* 2009;106: 14728-33.
- [30] Holmes E, Nicholls AW, Lindon JC, et al. Development of a model for classification of toxin-induced lesions using ¹H NMR spectroscopy of urine combined with pattern recognition. *NMR Biomed* 1998;11: 235-44.

- [31] Aharoni A, Ric d, V, Verhoeven HA, et al. Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry. *OMICS* 2002;6: 217-34.
- [32] Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass Spectrom Rev* 2007;26: 51-78.
- [33] Fiehn O, Kopka J, Trethewey RN, Willmitzer L. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal Chem* 2000;72: 3573-80.
- [34] Huang G, Chen H, Zhang X, Cooks RG, Ouyang Z. Rapid screening of anabolic steroids in urine by reactive Desorption Electrospray ionization. *Analytical Chemistry* 2007;79: 8327-32.
- [35] Jia LW, Wang C, Zhao SM, Lu X, Xu GW. Metabolomic identification of potential phospholipid biomarkers for chronic glomerulonephritis by using high performance liquid chromatography-mass spectrometry. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* 2007;860: 134-40.
- [36] Trauger SA, Go EP, Shen Z, et al. High sensitivity and analyte capture with desorption/ionization mass spectrometry on silylated porous silicon. *Anal Chem* 2004;76: 4484-9.
- [37] Vaidyanathan S, Gaskell S, Goodacre R. Matrix-suppressed laser desorption/ionisation mass spectrometry and its suitability for metabolome analyses. *Rapid Commun Mass Spectrom* 2006;20: 1192-8.
- [38] Vaidyanathan S, Jones D, Ellis J, et al. Laser desorption/ionization mass spectrometry on porous silicon for metabolome analyses: influence of surface oxidation. *Rapid Commun Mass Spectrom* 2007;21: 2157-66.
- [39] Wilson ID, Nicholson JK, Castro-Perez J, et al. High resolution "ultra performance" liquid chromatography coupled to oa-TOF mass spectrometry as a tool for differential metabolic pathway profiling in functional genomic studies. *J Proteome Res* 2005;4: 591-8.
- [40] Vaidyanathan S, Jones D, Jenkins T, Kell DB, Goodacre R. Metabolomic investigations using laser desorption ionisation mass spectrometry on porous silicon. *Abstracts of Papers of the American Chemical Society* 2005;229: U245.
- [41] Lindon JC, Nicholson JK. Spectroscopic and statistical techniques for information recovery in metabonomics and metabolomics. *Annu Rev Anal Chem* 2008;1: 45-69.
- [42] Lenz EM, Bright J, Wilson ID, Morgan SR, Nash AF. A ¹H NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects. *J Pharm Biomed Anal* 2003;33: 1103-15.
- [43] Nicholson JK, Higham DP, Timbrell JA, Sadler PJ. Quantitative high resolution ¹H NMR urinalysis studies on the biochemical effects of cadmium in the rat. *Mol Pharmacol* 1989;36: 398-404.

- [44] Nicholson JK, Connelly J, Lindon JC, Holmes E. Metabonomics: a platform for studying drug toxicity and gene function. *Nat Rev Drug Discov* 2002;1: 153-61.
- [45] Park SY, Kim YW, Kim JE, Kim JY. Age-associated changes in fat metabolism in the rat and its relation to sympathetic activity. *Life Sci* 2006;79: 2228-33.
- [46] Hodson MP, Dear GJ, Roberts AD, et al. A gender-specific discriminator in Sprague-Dawley rat urine: the deployment of a metabolic profiling strategy for biomarker discovery and identification. *Anal Biochem* 2007;362: 182-92.
- [47] Plumb R, Granger J, Stumpf C, et al. Metabonomic analysis of mouse urine by liquid-chromatography-time of flight mass spectrometry (LC-TOFMS): detection of strain, diurnal and gender differences. *Analyst* 2003;128: 819-23.
- [48] Williams RE, Lenz EM, Evans JA, et al. A combined (1)H NMR and HPLC-MS-based metabonomic study of urine from obese (fa/fa) Zucker and normal Wistar-derived rats. *J Pharm Biomed Anal* 2005;38: 465-71.
- [49] Ebbels TM, Holmes E, Lindon JC, Nicholson JK. Evaluation of metabolic variation in normal rat strains from a statistical analysis of 1H NMR spectra of urine. *J Pharm Biomed Anal* 2004;36: 823-33.
- [50] Gavaghan CL, Holmes E, Lenz E, Wilson ID, Nicholson JK. An NMR-based metabonomic approach to investigate the biochemical consequences of genetic strain differences: application to the C57BL10J and Alpk:ApfCD mouse. *FEBS Lett* 2000;484: 169-74.
- [51] Gu H, Chen H, Pan Z, et al. Monitoring diet effects via biofluids and their implications for metabolomics studies. *Anal Chem* 2007;79: 89-97.
- [52] Phipps AN, Stewart J, Wright B, Wilson ID. Effect of diet on the urinary excretion of hippuric acid and other dietary-derived aromatics in rat. A complex interaction between diet, gut microflora and substrate specificity. *Xenobiotica* 1998;28: 527-37.
- [53] Wikoff WR, Anfora AT, Liu J, et al. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci U S A* 2009;106: 3698-703.
- [54] Lawton KA, Berger A, Mitchell M, et al. Analysis of the adult human plasma metabolome. *Pharmacogenomics* 2008;9: 383-97.
- [55] Saude EJ, Adamko D, Rowe BH, Marrie T, Sykes BD. Variation of metabolites in normal human urine. *Metabolomics* 2007;3: 439-51.
- [56] Warrack BM, Hnatyshyn S, Ott KH, et al. Normalization strategies for metabonomic analysis of urine samples. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009;877: 547-52.
- [57] Jankevics A, Liepinsh E, Liepinsh E, et al. Metabolomic studies of experimental diabetic urine samples by H-1 NMR spectroscopy and LC/MS method. *Chemometrics and Intelligent Laboratory Systems* 2009;97: 11-7.

- [58] Wang J, Reijmers T, Chen L, et al. Systems toxicology study of doxorubicin on rats using ultra performance liquid chromatography coupled with mass spectrometry based metabolomics. *Metabolomics* 2009;5: 407-18.
- [59] Werner E, Croixmarie V, Umbdenstock T, et al. Mass spectrometry-based metabolomics: accelerating the characterization of discriminating signals by combining statistical correlations and ultrahigh resolution. *Anal Chem* 2008;80: 4918-32.
- [60] Bruce SJ, Tavazzi I, Parisod V, et al. Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. *Anal Chem* 2009;81: 3285-96.
- [61] Want EJ, O'Maille G, Smith CA, et al. Solvent-dependent metabolite distribution, clustering, and protein extraction for serum profiling with mass spectrometry. *Anal Chem* 2006;78: 743-52.
- [62] Croixmarie V, Umbdenstock T, Cloarec O, et al. Integrated comparison of drug-related and drug-induced ultra performance liquid chromatography/mass spectrometry metabolomic profiles using human hepatocyte cultures. *Anal Chem* 2009;81: 6061-9.
- [63] Soga T, Baran R, Suematsu M, et al. Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption. *J Biol Chem* 2006;281: 16768-76.
- [64] Soga T, Igarashi K, Ito C, et al. Metabolomic profiling of anionic metabolites by capillary electrophoresis mass spectrometry. *Anal Chem* 2009;81: 6165-74.
- [65] Idborg-Bjorkman H, Edlund PO, Kvalheim OM, Schuppe-Koistinen I, Jacobsson SP. Screening of biomarkers in rat urine using LC/electrospray ionization-MS and two-way data analysis. *Anal Chem* 2003;75: 4784-92.
- [66] Plumb RS, Stumpf CL, Gorenstein MV, et al. Metabonomics: the use of electrospray mass spectrometry coupled to reversed-phase liquid chromatography shows potential for the screening of rat urine in drug development. *Rapid Commun Mass Spectrom* 2002;16: 1991-6.
- [67] Lafaye A, Junot C, Ramounet-Le GB, et al. Metabolite profiling in rat urine by liquid chromatography/electrospray ion trap mass spectrometry. Application to the study of heavy metal toxicity. *Rapid Commun Mass Spectrom* 2003;17: 2541-9.
- [68] Madalinski G, Godat E, Alves S, et al. Direct introduction of biological samples into a LTQ-Orbitrap hybrid mass spectrometer as a tool for fast metabolome analysis. *Anal Chem* 2008;80: 3291-303.
- [69] Hodson MP, Dear GJ, Griffin JL, Haselden JN. An approach for the development and selection of chromatographic methods for high-throughput metabolomic screening of urine by ultra pressure LC-ESI-ToF-MS. *Metabolomics* 2009;5: 166-82.
- [70] Guy PA, Tavazzi I, Bruce SJ, Ramadan Z, Kochhar S. Global metabolic profiling analysis on human urine by UPLC-TOFMS: issues and method validation in

- nutritional metabolomics. *J Chromatogr B Analyt Technol Biomed Life Sci* 2008;871: 253-60.
- [71] Redestig H, Fukushima A, Stenlund H, et al. Compensation for Systematic Cross-Contribution Improves Normalization of Mass Spectrometry Based Metabolomics Data. *Analytical Chemistry* 2009;81: 7974-80.
- [72] Draisma HHM, Reijmers TH, van der Kloet F, et al. Equating, or Correction for Between-Block Effects with Application to Body Fluid LC-MS and NMR Metabolomics Data Sets. *Analytical Chemistry* 2010;82: 1039-46.
- [73] Wagner S, Scholz K, Sieber M, Kellert M, Voelkel W. Tools in metabonomics: an integrated validation approach for LC-MS metabolic profiling of mercapturic acids in human urine. *Anal Chem* 2007;79: 2918-26.
- [74] Rew RK, G.P.Davis. NetCDF: An Interface for Scientific Data Access. *IEEE computer graphics and applications* 1990;10: 76-82.
- [75] Pedrioli PG, Eng JK, Hubley R, et al. A common open representation of mass spectrometry data and its application to proteomics research. *Nat Biotechnol* 2004;22: 1459-66.
- [76] Katajamaa M, Oresic M. Data processing for mass spectrometry-based metabolomics. *J Chromatogr A* 2007;1158: 318-28.
- [77] Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 2006;78: 779-87.
- [78] Tautenhahn R, Bottcher C, Neumann S. Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics* 2008;9: 504.
- [79] Katajamaa M, Miettinen J, Oresic M. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* 2006;22: 634-6.
- [80] Baran R, Kochi H, Saito N, et al. MathDAMP: a package for differential analysis of metabolite profiles. *BMC Bioinformatics* 2006;7: 530.
- [81] Wiklund S, Johansson E, Sjostrom L, et al. Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Anal Chem* 2008;80: 115-22.
- [82] Trygg J, Holmes E, Lundstedt T. Chemometrics in metabonomics. *Journal of Proteome Research* 2007;6: 469-79.
- [83] Madsen R, Lundstedt T, Trygg J. Chemometrics in metabolomics-A review in human disease diagnosis. *Analytica Chimica Acta* 2010;659: 23-33.
- [84] Matsuda F, Yonekura-Sakakibara K, Niida R, et al. MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. *Plant J* 2009;57: 555-77.

- [85] Werner E, Heilier JF, Ducruix C, et al. Mass spectrometry for the identification of the discriminating signals from metabolomics: current status and future trends. *J Chromatogr B Analyt Technol Biomed Life Sci* 2008;871: 143-63.
- [86] Sumner LW, Urbanczyk-Wochniak E, Broeckling CD. Metabolomics data analysis, visualization, and integration. *Methods Mol Biol* 2007;406: 409-36.
- [87] Topol EJ. Failing the public health--rofecoxib, Merck, and the FDA. *N Engl J Med* 2004;351: 1707-9.
- [88] Wathion N., Public statement on Acomplia (rimonabant). Withdrawal of the marketing authorisation in the European Union. European Medicines Agency, 2009.
- [89] Gant TW. Novel and future applications of microarrays in toxicological research. *Expert Opinion on Drug Metabolism & Toxicology* 2007;3: 599-608.
- [90] Merrick BA, Tomer KB. Toxicoproteomics: A parallel approach to identifying biomarkers. *Environmental Health Perspectives* 2003;111: A578-A579.
- [91] Griffin JL. Metabonomics: NMR spectroscopy and pattern recognition analysis of body fluids and tissues for characterisation of xenobiotic toxicity and disease diagnosis. *Curr Opin Chem Biol* 2003;7: 648-54.
- [92] Keun HC. Metabonomic modeling of drug toxicity. *Pharmacol Ther* 2006;109: 92-106.
- [93] Lindon JC, Holmes E, Bollard ME, Stanley EG, Nicholson JK. Metabonomics technologies and their applications in physiological monitoring, drug safety assessment and disease diagnosis. *Biomarkers* 2004;9: 1-31.
- [94] Robertson DG. Metabonomics in toxicology: a review. *Toxicol Sci* 2005;85: 809-22.
- [95] Robertson DG, Reily MD, Baker JD. Metabonomics in preclinical drug development. *Expert Opin Drug Metab Toxicol* 2005;1: 363-76.
- [96] Shockcor JP, Holmes E. Metabonomic applications in toxicity screening and disease diagnosis. *Curr Top Med Chem* 2002;2: 35-51.
- [97] Timbrell JA. *Principles of Biochemical Toxicology*. Informa Healthcare 1999;3rd Revised edition.
- [98] Sun J, Schnackenberg LK, Holland RD, et al. Metabonomics evaluation of urine from rats given acute and chronic doses of acetaminophen using NMR and UPLC/MS. *J Chromatogr B Analyt Technol Biomed Life Sci* 2008;871: 328-40.
- [99] Schnackenberg LK, Chen M, Sun J, et al. Evaluations of the trans-sulfuration pathway in multiple liver toxicity studies. *Toxicol Appl Pharmacol* 2009;235: 25-32.
- [100] Sun J, Schnackenberg LK, Beger RD. Studies of acetaminophen and metabolites in urine and their correlations with toxicity using metabolomics. *Drug Metab Lett* 2009;3: 130-6.

- [101] Chen C, Krausz KW, Shah YM, Idle JR, Gonzalez FJ. Serum metabolomics reveals irreversible inhibition of fatty acid beta-oxidation through the suppression of PPARalpha activation as a contributing mechanism of acetaminophen-induced hepatotoxicity. *Chem Res Toxicol* 2009;22: 699-707.
- [102] La S, Yoo HH, Kim DH. Liquid chromatography-mass spectrometric analysis of urinary metabolites and their pattern recognition for the prediction of drug-induced hepatotoxicity. *Chem Res Toxicol* 2005;18: 1887-96.
- [103] Boudonck KJ, Mitchell MW, Nemet L, et al. Discovery of metabolomics biomarkers for early detection of nephrotoxicity. *Toxicol Pathol* 2009;37: 280-92.
- [104] van Vliet E, Morath S, Eskes C, et al. A novel in vitro metabolomics approach for neurotoxicity testing, proof of principle for methyl mercury chloride and caffeine. *Neurotoxicology* 2008;29: 1-12.
- [105] Ganote CE, Peterson DR, Carone FA. The nature of D-serine--induced nephrotoxicity. *Am J Pathol* 1974;77: 269-82.
- [106] Williams RE, Jacobsen M, Lock EA. ¹H NMR pattern recognition and ³¹P NMR studies with d-Serine in rat urine and kidney, time- and dose-related metabolic effects. *Chem Res Toxicol* 2003;16: 1207-16.
- [107] Williams RE, Major H, Lock EA, Lenz EM, Wilson ID. D-Serine-induced nephrotoxicity: a HPLC-TOF/MS-based metabonomics approach. *Toxicology* 2005;207: 179-90.
- [108] Ringeissen S, Connor SC, Brown HR, et al. Potential urinary and plasma biomarkers of peroxisome proliferation in the rat: identification of N-methylnicotinamide and N-methyl-4-pyridone-3-carboxamide by ¹H nuclear magnetic resonance and high performance liquid chromatography. *Biomarkers* 2003;8: 240-71.
- [109] Williams RE, Lock EA. Sodium benzoate attenuates D-serine induced nephrotoxicity in the rat. *Toxicology* 2005;207: 35-48.
- [110] McBurney RN, Hines WM, Von Tungeln LS, et al. The liver toxicity biomarker study: phase I design and preliminary results. *Toxicol Pathol* 2009;37: 52-64.
- [111] Chace DH. Mass spectrometry in the clinical laboratory. *Chemical Reviews* 2001;101: 445-77.
- [112] FOLCH J, ASCOLI I, LEES M, MEATH JA, LeBARON N. Preparation of lipide extracts from brain tissue. *J Biol Chem* 1951;191: 833-41.
- [113] Sabatine MS, Liu E, Morrow DA, et al. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation* 2005;112: 3868-75.
- [114] Lewis GD, Wei R, Liu E, et al. Metabolite profiling of blood from individuals undergoing planned myocardial infarction reveals early markers of myocardial injury. *J Clin Invest* 2008;118: 3503-12.

- [115] Turer AT, Stevens RD, Bain JR, et al. Metabolomic profiling reveals distinct patterns of myocardial substrate use in humans with coronary artery disease or left ventricular dysfunction during surgical ischemia/reperfusion. *Circulation* 2009;119: 1736-46.
- [116] Shaham O, Slate NG, Goldberger O, et al. A plasma signature of human mitochondrial disease revealed through metabolic profiling of spent media from cultured muscle cells. *Proc Natl Acad Sci U S A* 2010;107: 1571-5.
- [117] Oresic M, Simell S, Sysi-Aho M, et al. Dysregulation of lipid and amino acid metabolism precedes islet autoimmunity in children who later progress to type 1 diabetes. *J Exp Med* 2008;205: 2975-84.
- [118] Sreekumar A, Poisson LM, Rajendiran TM, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 2009;457: 910-4.
- [119] Ritchie SA, Ahiahonu PW, Jayasinghe D, et al. Reduced levels of hydroxylated, polyunsaturated ultra long-chain fatty acids in the serum of colorectal cancer patients: implications for early screening and detection. *BMC Med* 2010;8: 13.
- [120] Wishart DS. Human Metabolome Database: completing the 'human parts list'. *Pharmacogenomics* 2007;8: 683-6.
- [121] Kanehisa M, Araki M, Goto S, et al. KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 2008;36: D480-D484.
- [122] Smith CA, O'Maille G, Want EJ, et al. METLIN: a metabolite mass spectral database. *Ther Drug Monit* 2005;27: 747-51.
- [123] Bogusz MJ, Maier RD, Kruger KD, et al. Poor reproducibility of in-source collisional atmospheric pressure ionization mass spectra of toxicologically relevant drugs. *J Chromatogr A* 1999;844: 409-18.
- [124] Cui Q, Lewis IA, Hegeman AD, et al. Metabolite identification via the Madison Metabolomics Consortium Database. *Nat Biotechnol* 2008;26: 162-4.
- [125] Palit M, Mallard G. Fragmentation energy index for universalization of fragmentation energy in ion trap mass spectrometers for the analysis of chemical weapon convention related chemicals by atmospheric pressure ionization-tandem mass spectrometry analysis. *Anal Chem* 2009;81: 2477-85.
- [126] Oberacher H, Pavlic M, Libiseller K, et al. On the inter-instrument and inter-laboratory transferability of a tandem mass spectral reference library: 1. Results of an Austrian multicenter study. *J Mass Spectrom* 2009;44: 485-93.
- [127] Oberacher H, Pavlic M, Libiseller K, et al. On the inter-instrument and the inter-laboratory transferability of a tandem mass spectral reference library: 2. Optimization and characterization of the search algorithm. *J Mass Spectrom* 2009;44: 494-502.
- [128] Chan W, Cai Z. Aristolochic acid induced changes in the metabolic profile of rat urine. *J Pharm Biomed Anal* 2008;46: 757-62.

- [129] Chan W, Lee KC, Liu N, et al. Liquid chromatography/mass spectrometry for metabonomics investigation of the biochemical effects induced by aristolochic acid in rats: the use of information-dependent acquisition for biomarker identification. *Rapid Commun Mass Spectrom* 2008;22: 873-80.
- [130] Lin Y, Si D, Zhang Z, Liu C. An integrated metabonomic method for profiling of metabolic changes in carbon tetrachloride induced rat urine. *Toxicology* 2009;256: 191-200.
- [131] Ohta T, Masutomi N, Tsutsui N, et al. Untargeted metabolomic profiling as an evaluative tool of fenofibrate-induced toxicology in Fischer 344 male rats. *Toxicol Pathol* 2009;37: 521-35.
- [132] Sieber M, Wagner S, Rached E, et al. Metabonomic study of ochratoxin a toxicity in rats after repeated administration: phenotypic anchoring enhances the ability for biomarker discovery. *Chem Res Toxicol* 2009;22: 1221-31.
- [133] Lenz EM, Bright J, Knight R, Wilson ID, Major H. A metabonomic investigation of the biochemical effects of mercuric chloride in the rat using ¹H NMR and HPLC-TOF/MS: time dependent changes in the urinary profile of endogenous metabolites as a result of nephrotoxicity. *Analyst* 2004;129: 535-41.
- [134] Sun J, Von Tungeln LS, Hines W, Beger RD. Identification of metabolite profiles of the catechol-O-methyl transferase inhibitor tolcapone in rat urine using LC/MS-based metabonomics analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009;877: 2557-65.
- [135] Wagner S, Scholz K, Donegan M, et al. Metabonomics and biomarker discovery: LC-MS metabolic profiling and constant neutral loss scanning combined with multivariate data analysis for mercapturic acid analysis. *Anal Chem* 2006;78: 1296-305.
- [136] Lenz EM, Bright J, Knight R, Wilson ID, Major H. Cyclosporin A-induced changes in endogenous metabolites in rat urine: a metabonomic investigation using high field ¹H NMR spectroscopy, HPLC-TOF/MS and chemometrics. *J Pharm Biomed Anal* 2004;35: 599-608.
- [137] Patterson AD, Li H, Eichler GS, et al. UPLC-ESI-TOFMS-based metabolomics and gene expression dynamics inspector self-organizing metabolomic maps as tools for understanding the cellular response to ionizing radiation. *Anal Chem* 2008;80: 665-74.
- [138] Yang L, Xiong A, He Y, et al. Bile acids metabonomic study on the CCl₄- and alpha-naphthylisothiocyanate-induced animal models: quantitative analysis of 22 bile acids by ultraperformance liquid chromatography-mass spectrometry. *Chem Res Toxicol* 2008;21: 2280-8.
- [139] Kind T, Tolstikov V, Fiehn O, Weiss RH. A comprehensive urinary metabolomic approach for identifying kidney cancer. *Anal Biochem* 2007;363: 185-95.
- [140] Kim K, Aronov P, Zakharkin SO, et al. Urine metabolomics analysis for kidney cancer detection and biomarker discovery. *Mol Cell Proteomics* 2009;8: 558-70.

- [141] Woo HM, Kim KM, Choi MH, et al. Mass spectrometry based metabolomic approaches in urinary biomarker study of women's cancers. *Clin Chim Acta* 2009;400: 63-9.
- [142] Cho SH, Choi MH, Lee WY, Chung BC. Evaluation of urinary nucleosides in breast cancer patients before and after tumor removal. *Clin Biochem* 2009;42: 540-3.
- [143] Cho SH, Jung BH, Lee SH, et al. Direct determination of nucleosides in the urine of patients with breast cancer using column-switching liquid chromatography-tandem mass spectrometry. *Biomed Chromatogr* 2006;20: 1229-36.
- [144] Frickenschmidt A, Frohlich H, Bullinger D, et al. Metabonomics in cancer diagnosis: mass spectrometry-based profiling of urinary nucleosides from breast cancer patients. *Biomarkers* 2008;13: 435-49.
- [145] Guan W, Zhou M, Hampton CY, et al. Ovarian cancer detection from metabolomic liquid chromatography/mass spectrometry data by support vector machines. *BMC Bioinformatics* 2009;10: 259.
- [146] Issaq HJ, Nativ O, Waybright T, et al. Detection of bladder cancer in human urine by metabolomic profiling using high performance liquid chromatography/mass spectrometry. *J Urol* 2008;179: 2422-6.
- [147] Ma YL, Qin HL, Liu WJ, et al. Ultra-High Performance Liquid Chromatography-Mass Spectrometry for the Metabolomic Analysis of Urine in Colorectal Cancer. *Dig Dis Sci* 2009.
- [148] Yan SK, Wei BJ, Lin ZY, et al. A metabonomic approach to the diagnosis of oral squamous cell carcinoma, oral lichen planus and oral leukoplakia. *Oral Oncol* 2008;44: 477-83.
- [149] Osl M, Dreiseitl S, Pfeifer B, et al. A new rule-based algorithm for identifying metabolic markers in prostate cancer using tandem mass spectrometry. *Bioinformatics* 2008;24: 2908-14.
- [150] Wang Z, Tang WH, Cho L, Brennan DM, Hazen SL. Targeted metabolomic evaluation of arginine methylation and cardiovascular risks: potential mechanisms beyond nitric oxide synthase inhibition. *Arterioscler Thromb Vasc Biol* 2009;29: 1383-91.
- [151] Lin H, Zhang J, Gao P. Silent myocardial ischemia is associated with altered plasma phospholipids. *J Clin Lab Anal* 2009;23: 45-50.
- [152] Wang C, Kong H, Guan Y, et al. Plasma phospholipid metabolic profiling and biomarkers of type 2 diabetes mellitus based on high-performance liquid chromatography/electrospray mass spectrometry and multivariate statistical analysis. *Anal Chem* 2005;77: 4108-16.
- [153] Yang J, Zhao X, Liu X, et al. High performance liquid chromatography-mass spectrometry for metabonomics: potential biomarkers for acute deterioration of liver function in chronic hepatitis B. *J Proteome Res* 2006;5: 554-61.

- [154] Mutch DM, Fuhrmann JC, Rein D, et al. Metabolite profiling identifies candidate markers reflecting the clinical adaptations associated with Roux-en-Y gastric bypass surgery. *PLoS One* 2009;4: e7905.
- [155] van G, V, Verhey E, Poelmann R, et al. Metabolomics (liver and blood profiling) in a mouse model in response to fasting: a study of hepatic steatosis. *Biochim Biophys Acta* 2007;1771: 1263-70.
- [156] Yin P, Zhao X, Li Q, et al. Metabonomics study of intestinal fistulas based on ultraperformance liquid chromatography coupled with Q-TOF mass spectrometry (UPLC/Q-TOF MS). *J Proteome Res* 2006;5: 2135-43.
- [157] Jansson J, Willing B, Lucio M, et al. Metabolomics reveals metabolic biomarkers of Crohn's disease. *PLoS One* 2009;4: e6386.
- [158] Duarte IF, Legido-Quigley C, Parker DA, et al. Identification of metabolites in human hepatic bile using 800 MHz ¹H NMR spectroscopy, HPLC-NMR/MS and UPLC-MS. *Mol Biosyst* 2009;5: 180-90.
- [159] Li Q, Zhang Q, Xu G, et al. Metabonomics study of intestinal transplantation using ultrahigh-performance liquid chromatography time-of-flight mass spectrometry. *Digestion* 2008;77: 122-30.
- [160] Greenberg N, Grassano A, Thambisetty M, Lovestone S, Legido-Quigley C. A proposed metabolic strategy for monitoring disease progression in Alzheimer's disease. *Electrophoresis* 2009;30: 1235-9.
- [161] Rozen S, Cudkowicz ME, Bogdanov M, et al. Metabolomic analysis and signatures in motor neuron disease. *Metabolomics* 2005;1: 101-8.
- [162] Marhuenda-Egea FC, Martinez-Sabater E, Gonsalvez-Alvarez R, et al. A crucial step in assisted reproduction technology: human embryo selection using metabolomic evaluation. *Fertil Steril* 2009.
- [163] Shah DK, Doyle LW, Anderson PJ, et al. Adverse neurodevelopment in preterm infants with postnatal sepsis or necrotizing enterocolitis is mediated by white matter abnormalities on magnetic resonance imaging at term. *J Pediatr* 2008;153: 170-5, 175.
- [164] Yin P, Mohemaiti P, Chen J, et al. Serum metabolic profiling of abnormal savda by liquid chromatography/mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 2008;871: 322-7.
- [165] Lankinen M, Schwab U, Gopalacharyulu PV, et al. Dietary carbohydrate modification alters serum metabolic profiles in individuals with the metabolic syndrome. *Nutr Metab Cardiovasc Dis* 2009.
- [166] Gao P, Lu C, Zhang F, et al. Integrated GC-MS and LC-MS plasma metabonomics analysis of ankylosing spondylitis. *Analyst* 2008;133: 1214-20.
- [167] Beger RD, Holland RD, Sun J, et al. Metabonomics of acute kidney injury in children after cardiac surgery. *Pediatr Nephrol* 2008;23: 977-84.

- [168] Wikoff WR, Gangoiti JA, Barshop BA, Siuzdak G. Metabolomics identifies perturbations in human disorders of propionate metabolism. *Clin Chem* 2007;53: 2169-76.

Table1: MS-based metabolomics applications in toxicology

Experimental conditions	Topic	Application	Biological medium	comment	Reference
HPLC(C18)/QTOF-MS	Mechanistic considerations	Aristolochic induced nephrotoxicity	Rat urine		Chan et Al. 2008 [128]
HPLC(C18)/QTOF-MS	Mechanistic considerations	Aristolochic induced nephrotoxicity	Rat urine and plasma		Chan et Al. 2008 [129]
HPLC(HILIC) / IT-MS	Mechanistic considerations	CCl ₄ induced hepatotoxicity	Rat urine		Lin et al. 2009 [130]
HPLC(C18)/IT-MS	Mechanistic considerations	Fenofibrate-induced hepatotoxicity	Rat urine and plasma	Additional GC/MS experiments	Ohta et al. 2009 [131]
HPLC(C18)/QTOF-MS	Mechanistic considerations	D-serine induced nephrotoxicity	Rat urine	Previous study by 1H and 31P NMR (Williams et al. 2003)	Williams et al. 2005 [48]
UPLC(C18)/QTOF	Mechanistic considerations	Acetaminophen-induced hepatotoxicity : role of PPAR α	Mouse serum		Chen et al. 2009 [101]
UPLC(C18)/IT-MS	Mechanistic considerations and biomarker discovery	Ochratoxin A induced nephrotoxicity	Rat urine	Additional GC/MS and 1H NMR experiments	Sieber et al. 2008 [132]
HPLC(C18)/QTOF-MS	Mechanistic considerations and biomarker discovery	mercuric chloride induced nephrotoxicity	Rat urine	Additional 1H NMR experiments	Lenz et al. 2004 [133]
UPLC(C18)/TOF-MS	Biomarker discovery (xenometabolome)	Acetaminophen hepatotoxicity	Rat urine	Additional 1H NMR experiments	Sun et al. 2009 [100]
UPLC(C18)/TOF-MS	Biomarker discovery	Acetaminophen hepatotoxicity	Rat urine and serum	Additional 1H NMR experiments	Sun et al. 2008 [98]
UPLC(C18)/QTOF-MS	Biomarker discovery	Doxorubicine toxicity	Rat urine		Wang et al. 2009 [58]

UPLC(C18)/QTOF-MS	Biomarker discovery	Tolcapone toxicity	Rat urine		Sun et al. 2009 [134]
HPLC/LTQ-Orbitrap-MS	Biomarker discovery	Liver Toxicity Biomarker	Rat urine, plasma and liver	Targeted: lipid, AA and polar LC-MS metabolomics	McBurney et al. 2009 [110]
UPLC(C18)/TQ-MS	Biomarker discovery	Drugs-Induced Hepatotoxicity	Rat urine	Additional 1H NMR experiments	Schnackenberg et al. 2009 [99]
HPLC(C18)/QTOF-MS and QTRAP-MS	Biomarker discovery	Detection of mercapturic acid metabolites of acetaminophen	Human urine		Wagner et al. 2006 [135]
HPLC(C18)/QTRAP-MS	Biomarker discovery	Detection of mercapturic acid	Human urine		Wagner et al. 2007 [73]
HPLC(C18)/QTOF-MS	Biomarker discovery	Cyclosporine A induced nephrotoxicity	Rat urine	Additional 1H NMR experiments	Lenz et al. 2004 [136]
UPLC(C18)/QTOF-MS	Biomarker discovery	Ionizing radiation	Human cells		Patterson et al. 2008 [137]
UPLC/TQ-MS	Predictive model and biomarker discovery	CCl ₄ and ANIT induced hepatotoxicity	Rat serum	Targeted on bile acids	Yang et al. 2008 [138]
HPLC(C18)/IT-MS	Predictive model	Drug-induced nephrotoxicity	Rat urine and kidney	Additional GC/MS experiments	Boudonck et al. 2009 [103]
direct introduction/TQ-MS	Predictive model	Methyl mercury chloride induced neurotoxicity	Rat brain cell cultures		Van Vliet et al. 2008 [104]
HPLC(C18)/TQ-MS	Predictive model	Drug-induced hepatotoxicity	Rat urine		La et al. 2005 [102]

Table 2: MS-based metabolomics applications in clinical chemistry

Experimental conditions	Topic	Application	Biological medium	comment	Reference
UPLC(HILIC, C18)/IT-MS	Cancer	Kidney cancer	Human urine	Additional GC/MS experiments	Kind et al. 2007 [139]
HPLC(HILIC)/IT-MS	Cancer	Kidney cancer	Human urine		Kim et al. 2009 [140]
	Cancer	Woman's cancer (breast, ovarian, cervical)	Human urine	Targeted on hormones and nucleosides. Additional GC/MS experiments	Woo et al. 2009 [141]
HPLC(C18)/IT-MS	Cancer	Breast cancer before and after tumor resection	Human urine	Targeted on nucleosides	Cho et al. 2009 [142]
HPLC(C18)/IT-MS	Cancer	Breast cancer	Human urine	Targeted on nucleosides	Cho et al. 2006 [143]
HPLC(C18)/IT-MS HPLC/IT-MS	Cancer	Breast cancer	Human urine	Targeted on nucleosides	Frickenschmidt et al. 2008 [144]
HPLC(C18)/TOF-MS	Cancer	Ovarian cancer	Human serum		Guan et al. 2009 [145]
HPLC/IT-FT-MS	Cancer	Prostate cancer	Human plasma, urine and tissue	Additional GC/MS experiments	Sreekumar et al. 2009 [118]
HPLC(C18)/QTOF-MS or TQ-MS, and Direct introduction/FTICR-MS	Cancer	Colorectal cancer	Human serum	Both targeted (Q-TOF TQ-MRM) and non targeted (FTICR)	Ritchie et al. 2010 [119]
HPLC(C18)/QTOF-MS	Cancer	Bladder Cancer	Human urine		Issaq et al. 2008 [146]
UPLC(C18)/TOF-MS	Cancer	Colorectal cancer	Human urine		Ma et al. 2009 [147]
HPLC(C18)/TOF-MS	Cancer	Oral squamous cell carcinoma, oral lichen planus and oral leukoplakia	Human saliva		Yan et al. 2008 [148]
Direct introduction or HPLC/QTRAP-MS	Cancer	Prostate cancer	Human serum		Osl et al. 2008 [149]

HPLC/TQ-MS	Cardiology	Myocardial ischemia	Human plasma	Targeted: Sugars and ribonucleotides (luna normal phase column), Organic acids (polar-RP column) and amino acids (Luna phenyl-hexyl column)	Sabatine et al. 2010 [113]
HPLC/TQ-MS	Cardiology	Myocardial injury	Human plasma	Targeted: Sugars and ribonucleotides (luna normal phase column), Organic acids (polar-RP column) and amino acids (Luna phenyl-hexyl column)	Lewis et al. 2008 [114]
Direct introduction/TQ-MS	Cardiology	Coronary Artery Disease or Left Ventricular Dysfunction	Huma plasma	Targeted	Turer et al. 2009 [115]
HPLC(phenyl)/TQ-MS	Cardiology	Coronary artery disease	Human urine	Targeted on amino acids	Wang et al. 2009 [150]
UPLC(C18)/QTOF-MS	Cardiology	Silent myocardial ischemia	Human plasma		Lin et al. 2009 [151]
UPLC(HILIC)/TQ-MS	Endrocrinology	Type 1 diabetes	Rat urine	Additional 1H NMR experiments	Jankevics et al. 2009 [57]
UPLC(C18)/QTOF-MS	Endrocrinology	Type 1 diabetes	Human plasma	Targeted: lipidomics. Additional GC/MS experiments	Oresic et al. 2008 [117]
HPLC(nucleosil)/QTRAP-MS	Endrocrinology	Type 2 diabetes mellitus	Human plasma	Targeted on phospholipids	Wang et al. 2005 [152]
HPLC(C8)/QTRAP-MS	Hepato-Gastroenterology	Chronic hepatitis B	Human serum		Yang et al. 2006 [153]
HPLC(RP)/QTRAP-MS	Hepato-Gastroenterology	Roux-en-Y gastric bypass (RYGB) surgery	Human serum	Additional GC/MS experiments	Mutch et al. 2009 [154]
HPLC(C18)/IT-MS	Hepato-Gastroenterology	Hepatic steatosis	Mouse liver and blood	Targeted: lipidomics	Van Ginneken et al. 2007 [155]
UPLC(C18)/QTOF-MS	Hepato-Gastroenterology	Intestinal fistulas	Human blood		Yin et al. 2006 [156]
Direct introduction/FT-ICR-MS	Hepato-Gastroenterology	Crohn disease	Human fecal samples		Jansson et al. 2009 [157]

HPLC(C18)/IT-MS or UPLC(C18)/TOF-MS	Hepato-Gastroenterology	Liver transplantation	Human whole bile and extract	Additional 1H NMR and HPLC-NMR/MS experiments	Duarte et al. 2009 [158]
UPLC(C18)/QTOF-MS	Hepato-Gastroenterology	Intestinal transplantation	Human plasma		Li et al. 2008 [159]
UPLC(C18)/QTOF-MS	Neurology	Alzheimer's disease	Human plasma		Greenberg et al. 2009 [160]
HPLC(C18)/QTOF-MS or IT-MS	Neurology	Motor neuron disease	Human plasma	Additional GC/MS experiments	Rozen et al. 2005 [161]
HPLC(Luna SCX)/IT-MS	Reproduction	Human embryo selection	Human embryo culture medium		Marhuenda-Egea et al. 2009 [162]
HPLC(C18)/IT-MS	Reproduction	Preterm birth biomarkers	Human cervicovaginal fluid		Shah et al. 2008 [163]
UPLC(C18)/QTOF-MS	Uigur chinese traditional medicine	abnormal savda	Human serum		Yin et al. 2008 [164]
UPLC(C18)/QTOF-MS	Endocrinology and metabolism	Dietary carbohydrate and metabolic syndrome	Human plasma and tissue	Targeted: lipidomics. Additional GC/MS experiments	Lankinen et al. 2009 [165]
UPLC(C18)/QTOF-MS	Rheumatology	Ankylosing spondylitis	Human plasma	Additional GC/MS experiments	Gao et al. 2008 [166]
UPLC(C18)/QTOF-MS	Nephrology	Kidney injury in children after cardiac surgery	Human urine		Beger et al. 2007 [167]
HPLC(C18)/TOF-MS	Inborn errors of metabolism	Methylmalonic acidemia (MMA) and propionic acidemia (PA)	Human plasma		Wikoff et al. 2007 [168]
HPLC/QTRAP-MS	Inborn errors of metabolism	Respiratory chain diseases	Human plasma	Additional 1H NMR experiments. Targeted: Sugars and ribonucleotides (Luna normal phase column), Organic acids (polar-RP or ion pairing column) and amino acids (Luna phenyl-hexyl or HILIC column)	Shaham et al. 2010 [116]



Figure Legends

Figure 1: Schematic representation of omics technologies. The flow of information starts from genes to metabolites running through transcripts and proteins.

Figure 2: Number of publications dealing with metabolomics based on LC/MS in the fields of toxicology, clinical chemistry and others in the past 10 years (left axis) versus all metabolomics related publications (right axis). Search criteria in Pubmed were (metabolomics OR metabonomics) AND (liquid chromatography) AND (mass spectrometry) AND subject AND year[DP]" (subject are toxicology or disease, year ranges from 1999 to 2009).

Figure 3: The MS-based metabolomics flow chart.

Figure 4: Multivariate statistical analyses

Multivariate statistical analyses results are summarized into score and loading plots. The score plot represents the projection in two dimensions of samples onto principal components (PCs). The PCs constitute a new space which best carries the variation in the original data. The score plot shows how samples are dispersed in a 2- or 3-dimension space. Samples belonging to the same group are close from each other. The loading plot represents the projection of variables (m/z and retention time) onto PCs. Variables responsible for the discrimination between groups are far from the center of the loading plot, as emphasized with the two bar plots.

Figure 5: How to address signal redundancy?

This figure represents a Liquid Chromatography (LC) - Mass Spectrometry (MS) process and shows the origin of redundancy in MS signals. Four molecules represented by orange, green, yellow and blue dots are separated by LC. At retention time $t=1$, the "yellow molecule" is introduced into the ESI source. Into the source, pseudo molecular, adducts (e.g. with formic acid) and fragments (e.g. loss of functional group) ions are formed during the ionisation process (A). The resulting mass spectrum (B) reports the presence of those ions. The isotopic pattern of each ion (e.g. those of pseudo-molecular ion (C)) could be visualized when

enlarging the scale around the ion. Signals appearing at M+1, M+2 correspond to the isotopologues (^{13}C , ^{15}N ...) of the ion. When fragments and adducts as well as isotopologue ions are taken account, it appears that many signals are actually related to the single yellow molecules. This phenomenon is called signal redundancy.

Figure 6:

Top of the figure: spermidine is analyzed by flow injection analysis–high resolution mass spectrometry (FIA-HRMS). The mass spectrum is exported as a list of signals (with composition and attribution) that constitutes a home-made spectral database after data interpretation.

Bottom of the figure: Samples are analyzed by UPLC-MS and ions are extracted using automatic signal detection software. The m/z–retention time list is firstly annotated by search in Kegg, HMDB and Metlin databases and secondly by search in home-made spectral database.