Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography B, 866 (2008) 64–76
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Review

LC–MS-based metabonomics analysis☆

Xin Lu, Xinjie Zhao, Changmin Bai, Chunxia Zhao, Guo Lu, Guowang Xu*

National Chromatographic R&A Center, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China

Received 27 July 2007; accepted 16 October 2007
Available online 22 October 2007

Abstract

Metabonomics aims at the comprehensive and quantitative analysis of wide arrays of metabolites in biological samples. It has shown particular promise in the areas of toxicology and drug development, functional genomics, systems biology, and clinical diagnosis. Comprehensive metabonomics investigations are primarily a challenge for analytical chemistry. High-performance liquid chromatography–mass spectrometry (HPLC–MS) is an established technology in drug metabolite analysis and is now expanding into endogenous metabolite research. Its main advantages include wide dynamic range, reproducible quantitative analysis, and the ability to analyze biofluids with extreme molecular complexity. The aims of developing HPLC–MS for metabonomics range from understanding basic biochemistry to biomarker discovery and the structural characterization of physiologically important metabolites. In this review, the strategy and application of HPLC–MS-based metabonomics are reviewed.
© 2007 Elsevier B.V. All rights reserved.

Keywords: HPLC–MS; Metabonomics; Metabolic profiling; Metabolomics

Contents

1. Introduction . . . 65
2. Data collection methods of LC–MS-based metabonomics . . . 66
   2.1. Conventional HPLC–MS . . . 66
   2.2. Capillary HPLC–MS . . . 66
   2.3. Ultra-performance liquid chromatography (UPLC)–MS . . . 67
   2.4. On-line multidimensional high-performance chromatography method for metabonomics . . . 70
   2.5. MS ionization techniques and analyzers . . . 70
3. Data analysis of LC–MS-based metabonomics . . . 72
   3.1. Data processing . . . 72
   3.2. Pattern recognition . . . 72
   3.3. Databases . . . 72
   3.4. Metabolite identification . . . 73
4. Selected applications . . . 73
   4.1. Pharmaceutical research and development . . . 73
   4.2. Diseases biomarker discovery . . . 74
5. Future prospects . . . 74
   Acknowledgments . . . 75
   References . . . 75

☆ This paper is part of a Special Issue dedicated to the 50th anniversary of Journal of Chromatography.

* Corresponding author. Tel.: +86 411 84379530; fax: +86 411 84379559.
E-mail address: xugw@dicp.ac.cn (G. Xu).

1570-0232/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2007.10.022

1. Introduction

Metabonomics is defined as 'a quantitative measurement of multi-parametric metabolic responses of multi-cellular systems to pathophysiological stimuli or genetic signaling' [1]. Metabonomics has recently begun to have a more important role in efforts towards biomarker discovery. Metabonomics is typically performed on biofluids, such as serum, urine, saliva and cerebrospinal fluid. It is useful for physiological evaluation, drug safety assessment, diagnosis of human disease, drug therapy monitoring and characterization of genetically modified animal models of disease [2,3]. In combination with genomics, transcriptomics and proteomics, metabonomic analysis is being increasingly used in the drug discovery and development. Much, therefore, depends on the ability of the analytical technique employed to detect often-subtle differences in the complex mixtures found for biofluids. An extensive literature exists on the use of metabonomics to evaluate nephrotoxicants and hepatic toxicants [4–6].

There are two closely related terms metabonomics and metabolomics. The term metabolomics has traditionally been reserved for the analysis of plant metabolites, often performed in pursuit of functional genomic assessment [7,8]. In contrast, metabonomics is largely associated with ^{1}H NMR combined with pattern recognition to follow changes in metabolite flux in mammalian matrices, such as biofluid. Nicholson and coworkers have offered more precise definition for these two terms [9], and the methods and approaches used in the two disciplines are now highly convergent. They shall be considered interchangeable in the context of this review.

Fig. 1 shows the scheme for metabonomics investigations. From the point of view of analytical chemistry, it contains two key steps, that is, data collection of metabolite fingerprints (metabonome) and information mining. The comprehensive investigation of the metabolome is being complicated by its enormous complexity and dynamics. An ideal analytical technique can be performed directly on the samples, without the need for troublesome sample pretreatment. It should be a high-throughput screening, unbiased with the whole metabolites, robust, reproducible, sensitive and accurate. In addition, it should have a wide dynamic range. In an ideal world, all of these desirable characteristics would also be combined with sufficiently high-information content to enable the key metabolites identified via the post analysis multivariate statistical analysis of the data. A number of different analytical strategies are employed, with the ultimate goal of measuring a large fraction or all of the metabolites present. Realistically, no technique is currently available that can provide all of the desired properties. The major analytical techniques that are employed for metabonomics investigations are based on nuclear magnetic resonance (NMR) spectroscopy [1,3,10] and mass spectrometry (MS) [11–14]. Conventionally, metabonomics investigations have been performed using high-field NMR spectroscopy [3,6,15–17]. In this type of application, NMR has many advantages, such as high-information content of the resulting spectra, the relative stability of NMR-chemical shifts, the ease of quantification and the lack of any need to pre-select the conditions employed for the analysis. MS, besides NMR spectroscopy, is the main analytical technique widely used for metabonomic studies. Its main features are high sensitivity, high resolution, wide dynamic range, coverage of a wide chemical diversity, robustness and feasibility to elucidate the molecular weight (MW) and structure of unknown compounds. MS is inherently considerably more sensitive than NMR spectroscopy [18], but hyphenated chromatography (gas chromatography (GC) after chemical derivatization or high-performance liquid chromatography (HPLC)) enables preseparation of metabolites. Recently, LC–MS-based analytical technique has begun to be employed in metabonomic studies [14], either alone [12,19,20], or in combination with NMR spectroscopy [21–23].

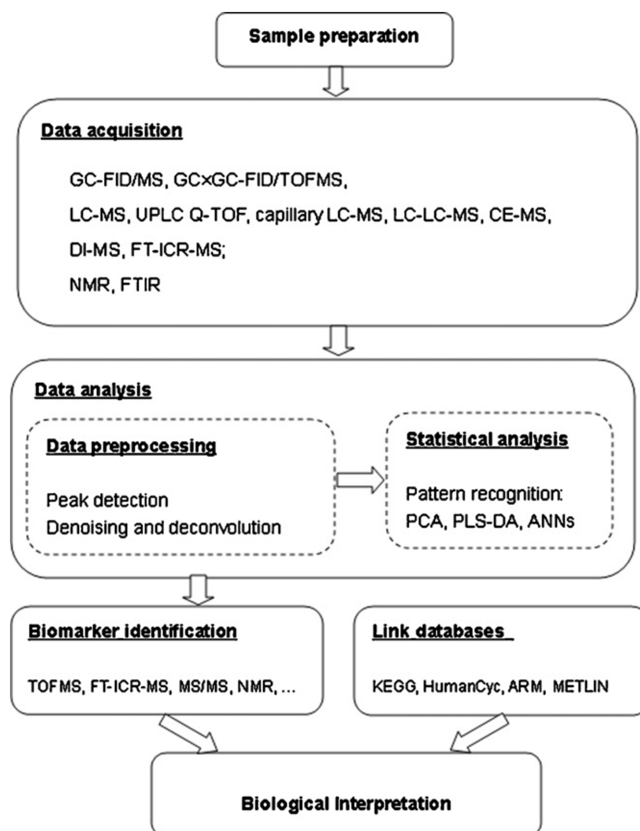

Fig. 1. Scheme for metabonomics investigations.

In information mining step, the identification of differences between samples has normally involved different multivariate tools such as principal component analysis (PCA), hierarchical cluster analysis (HCA), discriminate analysis or correlative network analysis. Ideally, the nonprocessed MS files would be subjected directly to multivariate analysis, but certain problems make preprocessing of raw data before multivariate analysis essential, including baseline problems, retention time drifts, variations in peak shape, and differences in recovery between the analyzed samples. In the meantime, in the identification step, because of lack of standard samples and database, an integrated strategy including various chromatography–MS techniques and NMR has to be used.

In this review, the main LC–MS-based technologies and data analysis methods used in metabonomics are summarized, and

the current applications of using LC–MS-based metabonomics are described. Some prospects for the future are then discussed.

2. Data collection methods of LC–MS-based metabonomics

2.1. Conventional HPLC–MS

The application of HPLC–MS for metabonomics studies is relatively new. However, the widespread availability of LC–MS has resulted in a rapid and continuing increase in the number of publications using the technique for metabonomics [14]. In many ways, LC–MS is ideal for metabolite profiling. Biofluids such as urine can be directly injected, whereas samples such as plasma need minimal pretreatment (protein precipitation). LC–MS is also capable of moderate to high throughput. It has a reasonable dynamic range combined with good potential for biomarker identification (based on the spectral data generated).

In general, metabonomics investigations by HPLC–MS have been performed using solvent gradients, on reversed-phase packing materials, 3.0 or 4.6 mm i.d. columns, of length between 5 and 25 cm in containing 3–5 μm packing materials [14]. A typical HPLC–MS total ion current chromatogram (positive ESI) for a rat urine sample obtained from a Zucker (fa/fa) obese rat is shown in Fig. 2 [24]. There are many ions present in these TICs (ca. 1600 in this case). Using a variety of statistical approaches (e.g., principal components analysis, PCA) to compare such metabolic profiles, it has proved to be possible to discover differences that enable animals of different strains, genders, ages, or subject to different treatments, to be distinguished from each other.

Many biofluids, particularly urine, contain a vast array of highly polar molecules that are not retained well on the reversed-phase chromatography. Normal phase techniques, which result in the elution of less polar molecules first and thus the retention of more polar molecules, require a different solvent system to that used by reversed-phase chromatography, typically containing non-aqueous mobile phase. One option for the HPLC–MS analysis of such polar compounds using the so-called "HILIC" (hydrophilic interaction chromatography) has been demonstrated for rat urine following solid-phase extraction (SPE) [25]. The unretained fractions in RP-HPLC were then analyzed on a "ZIC-HILIC" column using a solvent gradient over 15 min (followed by a further 4 min isocratic elution before returning to the starting conditions for re-equilibration). HILIC approaches combined with ESI-MS techniques have already been applied to the analysis of dichloroacetic acid in rat blood and tissues [26], and with APCI mass spectrometry for the determination of 5-fluorouracil in plasma and tissues [27].

2.2. Capillary HPLC–MS

Conventional HPLC–MS method has been shown to be capable of reasonable resolution and moderate throughput. However, higher resolution alternatives to conventional HPLC for complex metabolite analysis would be useful for metabonomics investigations. Capillary HPLC should provide improvements in metabonomic studies compared to conventional HPLC. It provides high-chromatographic resolution, high-peak capacity and increased signal to noise due to more concentrated peaks and reduced ion suppression. In addition, the marked reduction in the amount of sample required for capillary HPLC easily enables the analysis of very small samples (a few μL). Capillary HPLC has already been used in the plant metabolomics area where monolithic C18 bonded silica capillary columns (0.2 mm i.d. and 30–90 cm in length) were used with MS detection for *Arabidopsis thaliana* extracts [28]. Granger et al. [29] described the results of a comparison of both conventional and capillary HPLC columns, packed with the same chromatographic phase, for metabonomic analysis of urine from male and female Zucker obese (fa/fa) rats. The capillary HPLC–MS system chosen for analysis of the Zucker rat urine samples was based on a

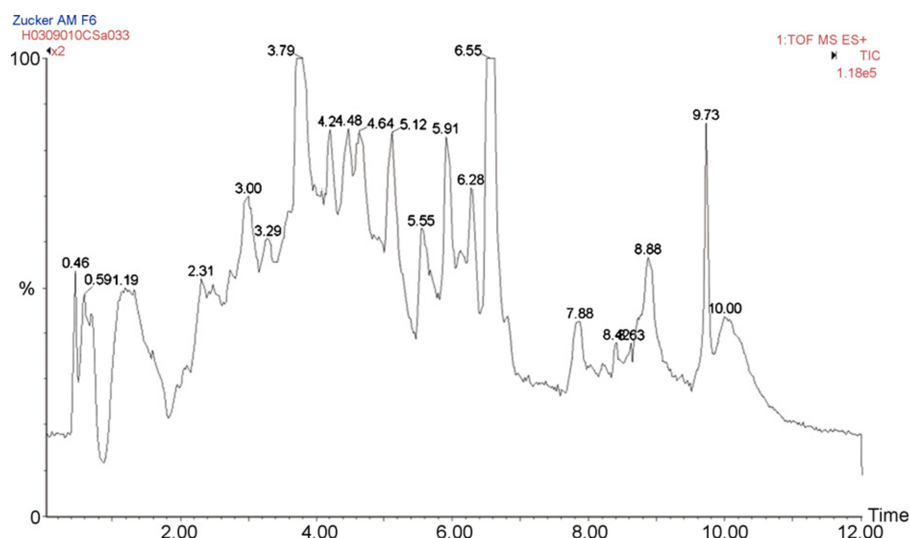

Fig. 2. Reserved-phase HPLC–MS, using positive ESI, of rat urine from a male Wistar-derived rat. Reproduced with permission from Lenz et al. [24].

320 μm capillary, 10 cm in length containing a 3.5 μm symmetry C18 alkyl bonded packing. Typical positive total ion current mass chromatograms and 2D mass *vs*. retention time plots for a male (AM) Zucker rat urine sample are shown in Fig. 3, obtained using conventional (Fig. 3a and b) and capillary (Fig. 3c and d) methods. From these results it can be seen that the overall pattern of peak distribution provided by both techniques is not dissimilar, with the bulk of the ions detected by both methods eluting between 2 and 5 min (Fig. 3c and d). However, the capillary method reveals the presence of many more components than the conventional separation. Capillary HPLC–MS provided increased sensitivity (ca. 100 folds) despite using a fraction of the sample volume consumed by HPLC (0.5 μL *vs*. 10 μL). Capillary HPLC–MS also provided a greater peak count (ca. 3000 ions compared to 1500 for HPLC) for the same analysis time compared to conventional HPLC–MS. This increased number of metabolites also enables increased discrimination between, e.g., samples obtained from female animals for a.m. and p.m. collection times. With the conventional HPLC–MS system there is a trend, but no clear separation of the a.m. and p.m. samples. In contrast, when such analysis is performed with the capillary there is a very clear separation. Capillary HPLC–MS has been shown to provide a superior means of obtaining metabolic profiles of urine than conventional HPLC–MS by detecting more peaks, with higher sensitivity and greatly reduced sample usage. The technique was sufficiently robust to enable the analysis of a large number of untreated urine samples without degradation of column performance. Capillary HPLC–MS may therefore offer considerable benefits to metabonomic/metabolomic analysis where sample volumes are limited.

2.3. Ultra-performance liquid chromatography (UPLC)–MS

The recently introduced UPLC is a combination of a 1.7 μm reversed-phase packing material and a chromatographic system, operating at pressures in the 6000–15000 psi range. Due to a reduction in band broadening, there will also be a greater S/N ratio, and thus an increase in sensitivity. This has enabled better chromatographic peak resolution and increased speed and sensitivity to be obtained for complex mixture separation. Reducing the particle diameter from 5 to 1.7 μm will, in principle, result in a 3-fold increase in efficiency and speed and a 1.7-fold increase in resolution and sensitivity. The typical peak widths generated by the UPLC system are in the order of 1–2 s for a 10 min separation. Because of the much improved chromatographic resolution of UPLC, the problem of ion suppression from co-eluting peaks is greatly reduced. Orthogonal-acceleration time-of-flight (oaTOF) MS, and a combination of quadrupole and oaTOF–MS (Q–TOF) coupled to UPLC, has proved to be a powerful tool for the identification of trace constituents of complex mixtures and/or for confirming their presence. Such instruments enable accurate mass measurement with accuracies of <5 ppm, which dispel interpretation ambiguities.

There have been an increasing number of applications of UPLC–MS, in the analysis of biological fluids in the field of metabonomics [13,19,30–33]. Wilson et al. [13] compared UPLC–MS with conventional HPLC–MS under similar analytical conditions. UPLC–MS method showed improved phenotypic classification capability and increased ability to probe differential pathway activities between strains as a result of improved analytical sensitivity and resolution. Fig. 4 shows

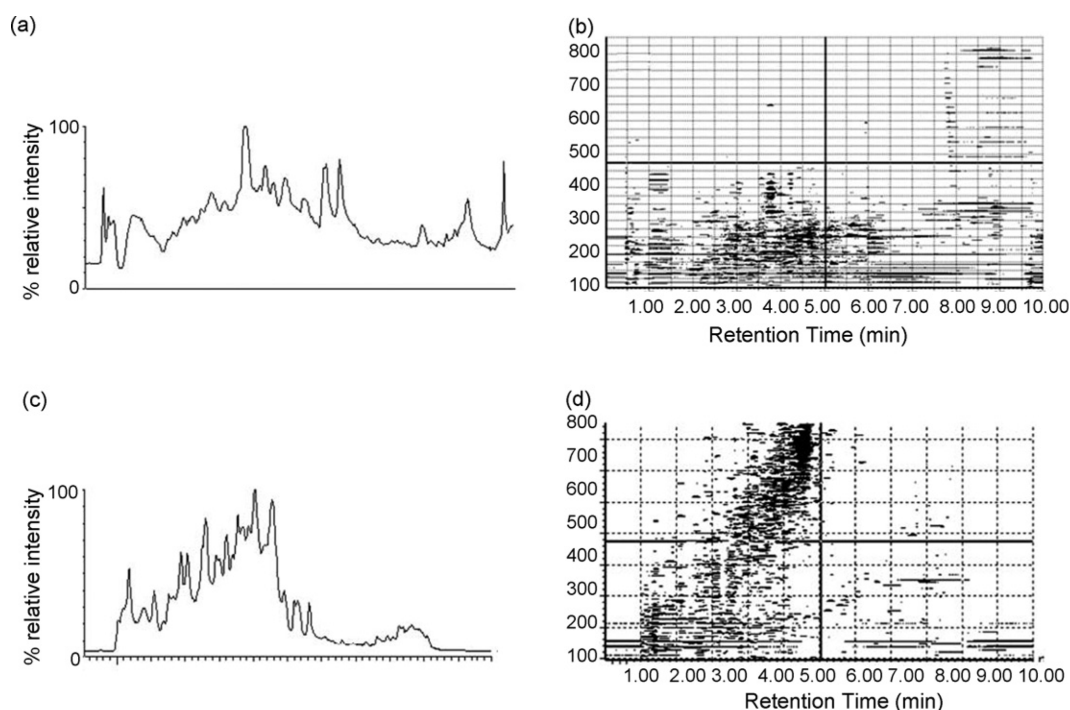

Fig. 3. Typical total ion current mass chromatogram, in positive ion ESI, for (a) HPLC and (c) capillary LC–MS with the corresponding two-dimensional ion maps (b) HPLC and (d) capillary LC for comparison. Reproduced with permission from Granger et al. [29].

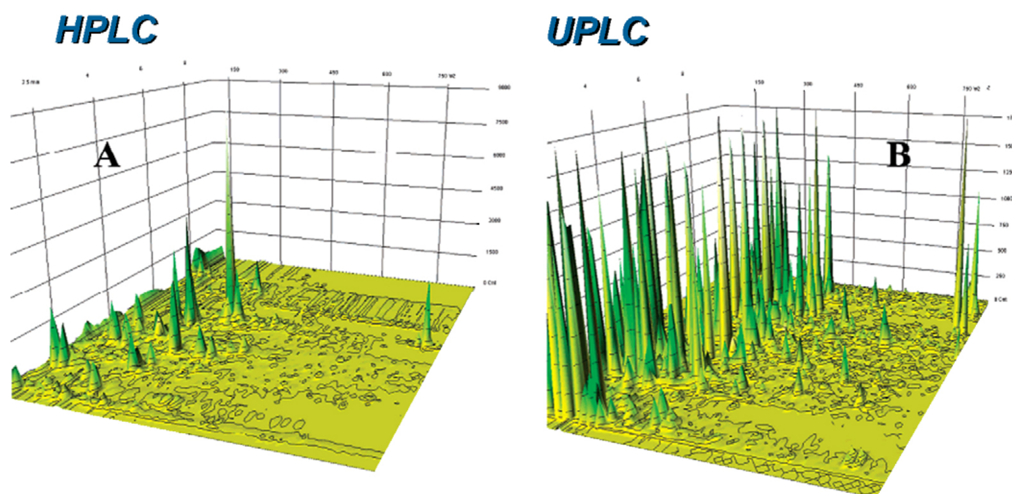

Fig. 4. Three-dimensional plots of retention time, *m*/*z*, and intensity from control white male mouse urine using (A) high-performance liquid chromatography–mass spectrometry (HPLC–MS) with a 2.1 cm × 100 mm Waters Symmetry 3.5 μm C18 column, eluted with 0–95% linear gradient of water with 0.1% formic acid/acetonitrile with 0.1% formic acid over 10 min at a flow rate of 0.6 mL/min and (B) ultra-performance liquid chromatography–mass spectrometry (UPLC–MS) with 2.1 cm × 100 mm Waters ACQUITY 1.7 μm C18 column, eluted with the same solvents at a flow rate of 0.5 mL/min. In both cases, the column eluent was monitored by electrospray ionization orthogonal-acceleration time-of-flight-mass spectrometry from 50 to 850 *m*/*z* in positive ion mode. Reproduced with permission from Wilson et al. [13].

the HPLC–MS and UPLC–MS chromatograms from a white male mouse urine sample. Several studies were carried out using UPLC–MS for the analysis of urine for a metabonomic investigation of gender, strain and diurnal variation in genetically distinct animals [13].

We used a UPLC/Q-TOF MS-based metabonomics technique to investigate intestinal fistula [32]. Sera from 40 patients with intestinal fistula and 17 healthy volunteers were collected and analyzed. After the UPLC/Q-TOF analysis, the retention time and *m*/*z* data pair for each peak were obtained. Both positive and negative ion TIC chromatograms are given in Fig. 5. According to peak height data, about 4900 peaks of positive ions (ESI^+) and 3600 peaks of negative ions (ESI^-) were detected by Micromass MarkerLynx. Partial least squares discriminant analysis (PLS-DA) and coefficient of correlation analysis were used for potential biomarker selection and identification. Fig. 6 shows the score and loading plot of PLS-DA. In both ion scan modes, the patients and controls obviously can be separated. Components that played important roles in the separation were picked out according to the parameter VIP (Variable Importance in the Projection). After an independent test, variables without significant differences between the patients and the controls ($p > 0.05$) were eliminated. Nine potential biomarkers were identified. Glycochenodeoxycholic acid, glycodeoxycholic acid (GDCA), taurochenodexycholic acid, taurodeoxycholic acid, and two kinds of lysophosphatidyl choline (C16:0 and C18:2) were found with increased concentrations in the patients, and phenylalanine, tryptophan, and carnitine were found with decreased concentrations in the patients. The results suggested that a subclinical hepatic injury and abnormal metabolism of two essential amino

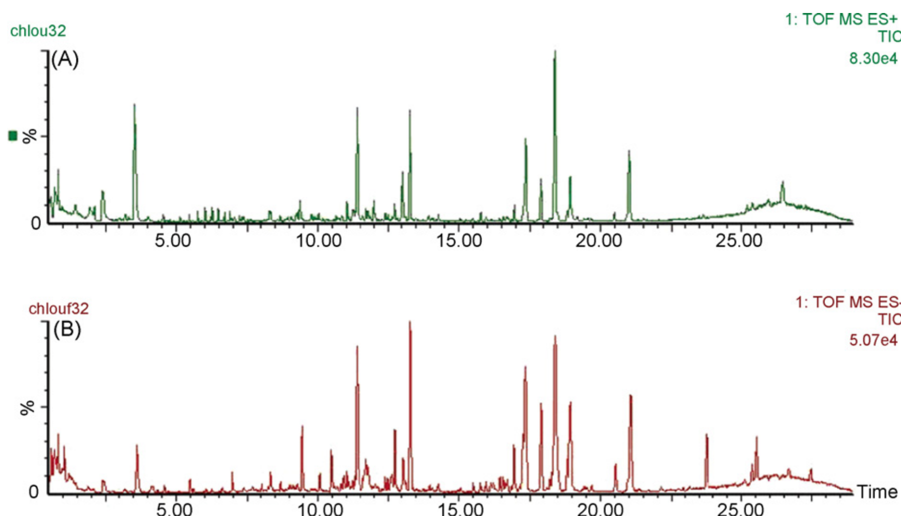

Fig. 5. Typical TIC chromatograms obtained from the same serum sample of a patient with the intestinal fistula. (A) ESI^+ scan; (B) ESI^- scan. Reproduced with permission from Yin et al. [32].

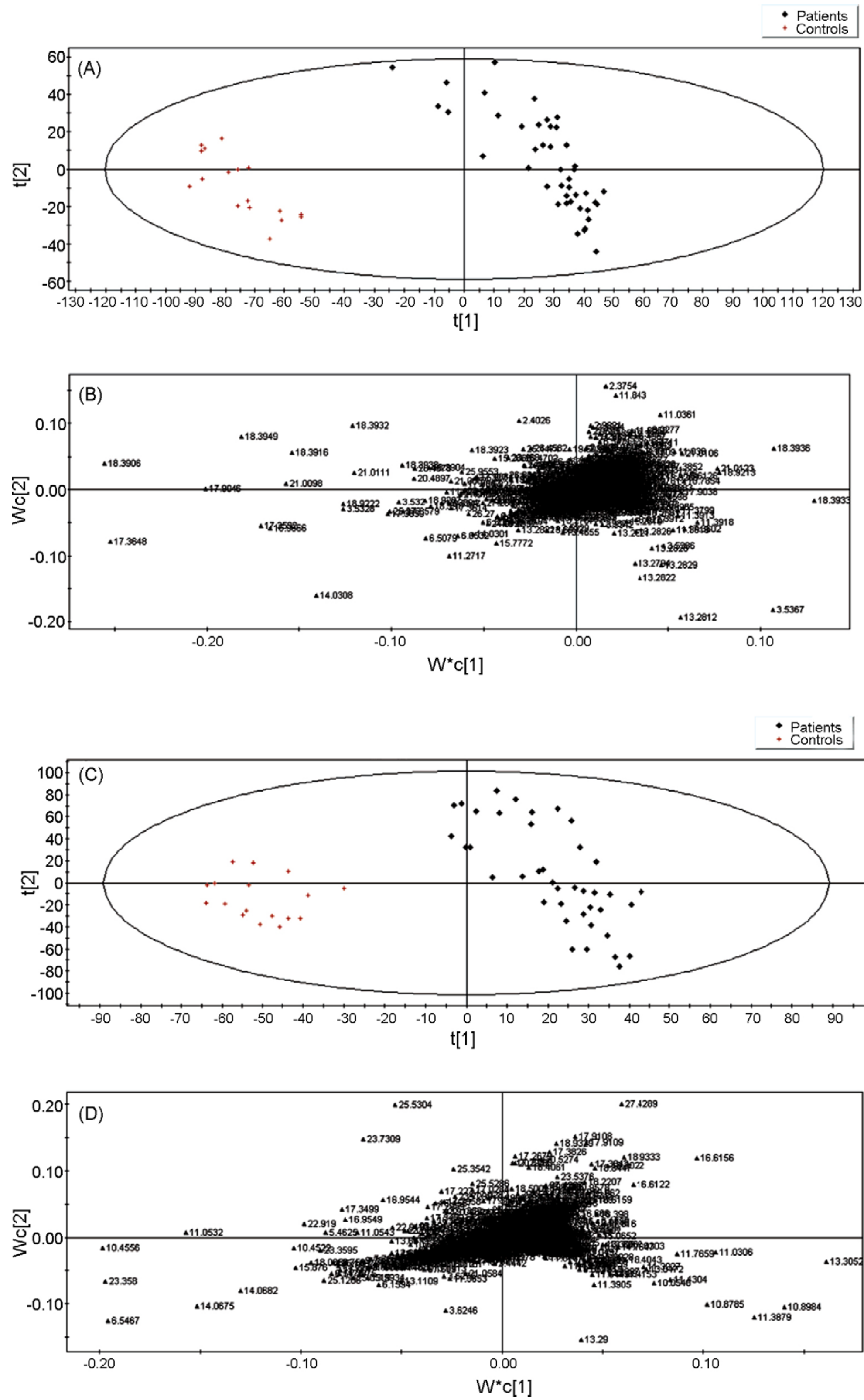

Fig. 6. PLS-DA result of the entire metabonome from 0.5 to 28 min. (A) PLS-DA score plot for the first two components indicating the separation between two groups, scanned by ESI$^+$; (B) PLS-DA variable loading plot explaining the separation above. The variables are labeled with retention time; (C) PLS-DA score plot, scanned by ESI$^-$; and (D) loading plot of panel C. (+) Controls. (⧫) Patients. Reproduced with permission from Yin et al. [32].

acids (phenylalanine and tryptophan), and a key compound of fatty acid synthesis and β-oxidation (carnitine), occurred in the fistula patients. In addition, we also developed an ultra-performance liquid chromatography (UPLC)-TOF MS method to analyze urinary nucleosides and other metabolites with *cis*-diol structure to distinguish cancer patients and healthy persons [33].

It should be emphasized that not only the UPLC with 1.7 μm packings from Waters, but also other HPLC with smaller diameters of packings, for example, RRLC with 1.8 μm packings from Agilent, Fast-HPLC with 2.2 μm from Shimadzu, will provide improved analytical characterization for metabolites.

2.4. On-line multidimensional high-performance chromatography method for metabonomics

Despite its high-resolution power, sensitivity, precision and practicability, HPLC analysis of biological samples like blood, serum, plasma, urine, milk, lymphatic fluids, liquor, feces or tissue homogenates is restricted by the pretreatment and processing of such highly complex matrices. The goal in bioanalytical sample processing should be a rapid and, if possible, an automated, HPLC integrated procedure that exhibits a high selectivity to the analyte and tolerates the direct injection of a biological sample. In many cases, metabolites are present in trace amounts and biological samples are too complex or incompatible with conventional HPLC phase systems to permit an analysis by direct injection into an analytical column. Thus, simplification of such multicomponent mixtures as well as metabolite enrichment is needed prior to analysis. In general, this is obtained by prefractionation or class separation and preconcentration steps.

To solve these problems, HPLC column-switching technique has been developed. Besides many other attractive applications in different analytical fields, this multidimensional separation method turned out as a powerful approach to enrich, separate and quantify a large variety of exogenous and endogenous compounds in complex biological samples such as blood, urine or tissue homogenates. In principle, these techniques always use two or more columns that are connected in parallel or in series and thus allow the selective prefractionation and subsequent analysis of the target compounds in deproteinized biological fluids. Boos et al. synthesized a series of precolumn packings and applied them in the coupled-column HPLC systems for the analysis of drugs and endogenous compounds in different biological matrices [34–36].

We developed a fully automated method for the determination of nucleoside metabolic profiling [37]. Efficient on-line clean-up and concentration of nucleosides from urine samples were obtained by using a boronic acid-substituted silica column (40 mm × 4.0 mm i.d.) as the first column and a Hypersil ODS2 column (250 mm × 4.6 mm i.d.) as the second column. Urinary nucleosides were analyzed in 50 min requiring only pH adjustment and the protein precipitation by centrifugation. The good linearity, precision, sensitivity and selectivity obtained with this method allow its use for metabolic profiling research.

In metabonomics studies, there has been a considerable interest in simultaneous obtaining both hydrophobic and hydrophilic metabolites. To address this requirement, in our laboratory, an automated column-switching HPLC separation system for simultaneous separation of hydrophilic and hydrophobic solutes in a complex sample has been established [38,39]. More recently, this column-switching system is being applied for metabonomics investigations [40], and the results will be reported in subsequent articles.

2.5. MS ionization techniques and analyzers

For metabonomics studies, electrospray ionization (ESI) is the most commonly used technique in LC–MS [41,42]. In order to obtain a comprehensive profile, ionization must be performed in both positive and/or negative mode. ESI utilizes a high-electric field to produce charged droplets from a liquid solution, ultimately leading to the formation of gas phase ions [43]. The main advantages of the ESI ion source are soft ionization, no need for derivatization, ability to ionize compounds of a large mass range, suitability for nonvolatile and polar compounds and excellent quantitative analysis and high sensitivity. A drawback of the ESI process is its liability for ion suppression due to competition effects in the ionization process.

In our previous study on intestinal fistula, we compared the effect of ESI^+ and ESI^- on potential biomarker discovery [32]. In both ion scan modes, the patients and controls can be separated obviously (Fig. 6). Components that played important roles in the separation were picked out according to the parameter VIP (Variable Importance in the Projection). As can be seen from Table 1, the markers selected from the ESI^+ and ESI^- scans were quite different, the combination of the ESI^+ and ESI^- scans provided us more useful information about the samples.

Three alternative solution-based ionization strategies to ESI are also being used for HPLC–MS-based metabonomics, namely nanoESI, atmospheric pressure chemical ionization (APCI) and atmospheric pressure photoionizaion (APPI). NanoESI liquid chromatography, performed at low-flow rates (∼200 nL/min), has already proved useful in proteomics studies, it significantly enhances sensitivity and dynamic range. Using nanoLC/nanoESI–MS methods, the chemical noise entering the mass spectrometer was decreased, and the chromatographic separation was improved. The application of capillary HPLC/nanoESI–MS methods [44] provides better separations and hence a smaller chance of co-elution with competitors for ionization. ESI–MS under nanospray conditions at sub μL/min flow rates may lead to more uniform response factors. Chip-based nanoelectrospray systems [45–47] are another tool to reduce matrix effects and improve sensitivity [48]. APCI [49] and APPI [50] are widely used in the pharmaceutical industry, yet they are not frequently used for metabomonics investigations. Analogous to the ESI interface, APCI and APPI typically induce little or no fragmentation and are considered robust and relatively tolerant of high-buffer concentration. It is now recognized that these approaches can be valuable for the analysis of nonpolar and thermally stable compounds such as lipids with the apparent trend toward a “single” ionization source containing combinations of ESI and APCI or ESI and APPI. Other important ionization techniques such as direct analysis in real

Table 1
Markers selected showed difference between ESI+ and ESI− scans when global serum metabonome data were used

t_r (min)	*m/z*	VIP[a]	*p*-Value[b]	Identification result
ESI+				
18.391	184.423	17.379	1E−9	Phosphatidylcholine moiety of LPC
17.365	184.422	17.116	3E−7	
17.904	184.423	13.637	1E−9	
17.359	502.362	11.578	5E−8	LPC C18:2 fragment
16.966	184.422	11.254	6E−7	Phosphatidylcholine moiety of LPC
18.39	478.393	10.562	1E−8	LPC C16:0 fragment
14.031	209.457	9.842	1E−6	UN
18.393	497.363	8.776	1E−11	Isotope of LPC C16:0
ESI−				
23.358	303.466	11.691	2E−4	UN
10.456	367.316	11.531	7E−4	UN
6.547	311.365	11.226	7E−9	Dipeptide of phenylalanine
11.053	369.333	8.937	3E−2	UN
13.305	227.422	8.934	1E−3	UN
14.067	223.438	8.756	3E−4	UN
10.899	528.196	7.919	8E−4	UN
14.068	267.393	7.527	8E−4	UN
11.387	464.336	6.733	0.01	GDCA
10.453	368.275	6.406	4E−2	UN

LPC: lysophosphatidyl choline GDCA: glycodeoxycholic acid UN: unidentified.
[a] Variable Importance in the Projection.
[b] *p*-Value of independent *t*-test.

time (DART), matrix laser desorption/ionization (MALDI) [51] are not frequently used for untargeted analyses of biofluids.

There are many types of mass analyzers available for interfacing with HPLC. The basic common mass analyzers include single-quadrupoles, triple-quadrupoles (TQ), time-of-flight (TOF), ion traps, orbitrap and fourier transform ion cyclotron resonance (FT-ICR). In addition to these types of mass analyzers, an increasing number of hybrid systems exist that combine two basic types of mass spectrometers, such as Q-TOF instruments [52], quadrupole linear ion traps (Q-Trap) [53,54], or ion trap FT mass spectrometers [55]. Quadrupole mass spectrometers are still the most widespread instruments today. Quardrupole instruments are robust, have a high-linear dynamic range and are capable of analyzing an *m/z* range of 50–4000. The single-quadrupole mass analyzer is the simplest option, and it provides only nominal mass resolution. Its low-duty cycle for full scan data acquisition also reduces the sensitivity and limits its application for metabolic profiling. Tandem mass spectrometry for structural investigations is possible with triple-quadrupole instruments only, whereas single-quadrupole systems are used as chromatographic mass detector. The triple-quadrupole mass spectrometer allowed not only *m/z* information, but also fragment ion information to be generated, allowing experiments such as neutral loss and product ion spectra to be performed. Using a process known as selected reaction monitoring, the first and third quadrupoles monitor the parent and product ion, respectively, of a fragmentation transition that is specific for the target analyte. Along with providing highly selective detection, this arrangement avoids the duty cycle limitation. Ion trap mass spectrometers are compact instruments that cover an *m/z* range of up to 6000 and are able to operate in full scan mode at high speed and at low resolution. One of the benefits of ion trap instruments is their capability of performing successive fragmentation steps (MS^n). The introduction of ion trap mass spectrometers allowed MS^n spectra to be produced, achieving more information on new drug metabolites to help confirm structures. This capability was exploited, e.g., by Dear et al. [56], who employed MS3 and MS4 to help characterize in vivo drug metabolite structures. Linear ion traps quadrupole hybrid instruments (Q Trap or QqLIT) combine the MS^n capabilities of ion trap instruments with the neutral loss and precursor ion scan capabilities of triple-quadrupole instruments. Therefore, the shortcomings of both approaches are overcome [57]. A drawback to conventional ion traps is that they do not possess accurate mass capability. TOF instruments feature fast scanning capabilities, wide mass range and high resolution (5000–20,000 full width, half maximum [FWHW]) and mass accuracy. It is extremely useful for profiling complex metabolic mixtures. In order to perform MS–MS experiments with TOF instruments, another mass analyzer has to be combined. In quadrupole-TOF instruments, the last quadrupole of the triple quad configuration is substituted by a TOF analyzer. These hybrid instruments combine the stability and robustness of the quadrupole analyzer with the fast scanning capabilities, accuracy (<5 ppm) and high sensitivity of TOF mass analyzers. Orbitrap [58,59] and FT-MS [60,61] offer the highest resolution available and high-accuracy fragment masses. The unsurpassed resolution (>100,000 FWHW) and mass accuracy (<2 ppm) of FT-MS leads to formula candidates and supports metabolite identification. For high-end hybrid instruments like quadrupole FT-MS and qudrupole linear ion trap FT-MS instruments, ion selection and fragmentation can be performed outside the cyclotron of the FT-MS. The performance and application depth of FT-MS is thereby significantly expanded, although quantitative aspects may be compromised due to the ion gating function of the ion trap.

Metabolic profiling approaches require a sensitive full scan mode and exact masses. Therefore, Q-TOF instruments or linear ion trap FT-MS instruments are advantageous. In contrast, for targeted analysis of selected metabolites, triple-quadrupole instruments, and especially Q-Trap instruments (with their capability for multiple reaction monitoring), are frequently used. The development of the hybrid quadrupole-orthogonal time-of-flight mass spectrometer allowed the generation of exact mass information with greater accuracy and precision; these mass values can then be used to produce candidate empirical formulae which, at the 3–5 ppm error range, significantly reduce the number of possible structures of putative metabolites with molecular masses of a few hundred Da.

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) is widely used in proteomics analysis, nucleotide sequencing, and polymer analysis [62–64]. It has the potential to contribute significantly to metabolomics. It offers the advantage of high throughput, the unique ability to generate singly charged ions of <1000 Da and its tolerance to moderate levels of salt. Matrix-

assisted laser desorption/ionization time-of-flight/time-of-flight (MALDI-TOF/TOF) tandem mass spectrometry additionally permits the discrimination of isomeric molecular species that would not be possible using spectra of molecular ions alone. Sun et al. [65] developed a shotgun metabolomics approach using MALDITOF/TOF MS for the rapid analysis of negatively charged water-soluble cellular metabolites. A total of 285 metabolites from murine myocardium were identified based on mass accuracy and 90 metabolites were confirmed by tandem MS analyses.

3. Data analysis of LC–MS-based metabonomics

3.1. Data processing

The analytical platform in metabonomics experiments generates large amounts of data from a single sample, which must be pre-processed before multivariate analysis. The reason is the two-dimensional nature (chromatogram/mass spectra) of the data for each sample. The data processing for HPLC–MS-based metabolic profiling usually proceeds through several stages. The extraction of relevant information and elimination of noise from full-scan metabolic profiling is the first step. Curve resolution or deconvolution methods are mainly applied for data processing that results in a multivariate profile for each sample. Decovolution can also be used to reduce the complexity of chromatograms obtained with soft ionization techniques by filtering multiple charged species, clusters, and adducts [66]. Normalization stage reduces the systematic error by adjusting the intensities within each sample run. For any statistical analysis it is a prerequisite that the data is aligned and variances between samples that are not attributed to true differences are reduced. For HPLC–MS, the retention time variation from chromatogram to chromatogram has been a significant impediment against the use of chemometrics techniques in the analysis of chromatographic data due to the variations in the mobile phase composition, gradient reproducibility, temperature variations and column variability, and drift in the *m/z* direction is fairly straightforward to be corrected by *m/z* calibration. Since a variable in a data table should define the same property over all samples, variability in retention time drift cause problems for statistical modeling. A multitude of different peak alignment methods have been developed. Several commercial and open source routines for automatic alignment, denoising, deconvolution and extraction of peak have been proposed [66–72], such as MS Resolver and ReOrder (Pattern Recognition Systems, Bergen, Norway) [72], MZmine [68,69], XCMS [67,70] and MET-IDEA [67]. Many instrument manufacturers have produced their own software, such as MarkerLynx (Waters), MassHunter (Agilent) and MarkerView (Applied Biosystems/MDS SCIEX). Then multivariate data analysis is used for interpretation of metabolic profiling data.

3.2. Pattern recognition

The purpose of data analysis in metabonomics is versatile, including identification of similarities and difference in data, classification of samples, identification and quantification of analytes and more. Both univariate and multivariate statistics follow the data reduction step, and both supervised and unsupervised strategies are employed for multivariate statistics and model building for prediction and classification of outcomes [73]. The most common chemometric tool used in the evaluation of a metabonomics study is principal components analysis (PCA). PCA is always recommended as starting point for analyzing multivariate data and will rapidly provide an overview of the information hidden in the data. PCA can be used to investigate clustering tendency, detect outliers, and to visualize data structure [74]. In addition, there are many other unsupervised methods, such as nonlinear mapping and hierarchical cluster analysis. PCA gives a simplified representation of the information contained in the spectra and cannot generally use additional information about data, such as class information. Therefore, PCA is often followed by a supervised analysis technique such as partial least squares discriminant analysis (PLS-DA). PLS is one of widely used supervised methods. It relates a data matrix containing independent variables from samples, such as spectral intensity values (an X matrix), to a matrix containing dependent variables (e.g., measurements of response, such as toxicity scores) for those samples (a Y matrix). PLS-DA is performed in order to enhance the separation between groups of observations. Orthogonal projection on latent structure discriminant analysis (O-PLS-DA), which is the most recent advanced development of PLS-DA, can improve the interpretation of models [75]. Methods from the field of artificial intelligence, in particular artificial neural networks (ANN), have been successfully applied to metabonomics. It is capable of learning patterns and relations from input data, making good pattern recognition engines and robust classifiers. ANNs are used for building nonlinear classification and regression models. Recently, an approach to the mining of highly complex metabolomics data is to apply evolutionary supervised learning techniques, including genetic algorithms, genetic programming, evolutionary programming and genomic computing, which could be ideal strategies for mining such high-dimensional data as that obtained from metabolomic studies [76]. An overview of how the underlying philosophy of chemometrics is integrated throughout metabonomics studies was given by Trygg et al. [77].

3.3. Databases

In contrast to the well-annotated gene and protein databases that can be searched easily, at present, no such comprehensive tools exist for metabolite researchers. However, current metabolite databases, although incomplete, offer a starting point for characterization. Biochemical databases can be used to identify unknown metabolites, for example, to identify structure from known elemental composition, or to determine the biological function of the identified metabolites. Among the databases currently available, the most widely used are the NIST database, which includes mass spectral data for some known metabolites [78], as well as the KEGG, HumanCyc, ARM, and METLIN databases. The KEGG database is a valuable resource for metabolomics researchers [79]. HumanCyc [80] includes

known metabolites as well as those predicted by algorithms which project metabolic pathways from a genomic sequence. A database constructed as part of the Atomic Reconstruction of Metabolism (ARM) project compiles metabolite structures together with molecular weight and MS fragmentation data [81]. The University of Alberta hosts a mini-library of full mass spectra of newer drugs, metabolites and some breakdown products [82]. The METLIN database [83] catalogues metabolites, MS/MS spectra, and LC–MS profiles of human plasma and urine samples [84]. The Spectral Database for Organic Compounds SDBS provides access to a wealth of spectra of organic compounds (NMR, MS, IR). Another metabolite database is the "tumor metabolome" database, established at the Justus-Liebig University Giessen in Germany [85]. More specific lipidomics databases exist, such as Lipid Maps [86], SphinGOMAP [87], and Lipid Bank [88], which contain structural and nomenclatural information as well as standard analytical protocols. General information about physico-chemical properties of metabolites can be obtained by searching general chemical databases such as PubChem or CAS.

To date, existing metabolomics databases aim primarily at the structural identification of metabolites in various biological samples. However, once a better annotation of the metabolome in various organisms is achieved, the generation of databases containing quantitative metabolite data can be expected. An example for this type of database is the human metabolite database that contains more than 1400 metabolites found in the human body. Each metabolite is described by a MetaboCard designed to contain chemical data, clinical data and molecular biology/biochemistry data [89,90].

3.4. Metabolite identification

Once potential feature has been identified from metabolic profiling investigation, the identification of the potential biomarker is required. Experimental approaches for the structure elucidation by mass spectrometry are extensively described in the literature [91–93]. Often a combination of different mass spectrometric techniques is required for the structural elucidation of unknowns. Database such as KEGG, human metabolite database, and METLIN can be used to search candidate molecules. Valuable information can be obtained by FT-ICR-MS because of its ultra-high resolution and mass accuracy, which allows derivation of molecular empirical formulae and its capability for high-resolution MS^n experiments. It also provides detailed information about molecular structure units. A less expensive alternative to FT-ICR-MS for MS^n experiments is a 3D ion trap mass analyzer. High-resolution Q-TOF instruments can be used for MS/MS experiments providing high-resolution data on fragment ions.

Directly coupled chromatography–NMR spectroscopy methods can also be used. The most general of these "hyphenated" approaches is HPLC–NMR–MS [94], in which the eluting HPLC peak is split, with parallel analysis by directly coupled NMR and MS techniques. This can be operated in on-flow, stopped-flow, and loop-storage modes and thus can provide the full array of NMR- and MS-based molecular identification tools.

In addition, metabolites can be identified by library search. GC–EI-MS spectra have been studied for several decades. The largest spectra libraries are available from Wiley (∼400,000 spectra) and the National Institute of Standards and Technology (NIST) (∼200,000 spectra). Spectral libraries for LC–MS are less developed than the corresponding GC–MS libraries. ESI and APCI spectra generally comprise pesudomolecular ions depending on the chemical properties of the analytes, nature of the matrix and solvent composition. The collisionally induced dissociation (CID) product ion mass spectrum provides fragment information. However, fragmentation spectra vary between different types of mass analyzers (triple quadrupole instruments and ion traps) and even between instruments of the same mass analyzer type but different brands [95,96]. Despite the described difficulties, MS/MS spectral libraries are being created. Recently NIST has published a new update of their library (NIST 2005) containing 5191 MS/MS spectra of 1943 different ions (1671 positive and 341 negative ions). The spectra were acquired with different triple quadrupole and ion trap instruments. To date 79 MS/MS spectra of common metabolites and drugs acquired with a Q-TOF instrument are available at the METLIN database [84]. To increase the specificity of library search, retention time is also included in the library. This approach was used for the identification of drug metabolites in urine samples [97]. A more detailed overview on currently available LC–MS and GC–MS libraries was recently published [98].

Despite the usefulness of mass spectra data, the lack of comprehensive mass spectral libraries often precludes identification of molecules based on this data alone. The combination of many technologies will be required to identify unknown metabolites in biofluids, including high-sensitivity capillary NMR which can provide metabolite structure characterization down to low-microgram level, chemical modification for functional group identification, and finally chemical synthesis of potential candidates for verification.

4. Selected applications

4.1. Pharmaceutical research and development

Mass spectrometry has been widely used for in metabolic fingerprinting and metabolite identification. Although most investigations have been on plant metabolism and model cell system extracts. Recently, its applications in mammalian studies are increasing [11,14,99]. The first reported rodent toxicology investigation of HPLC–MS-based metabonomics method was by Plumb and co-workers [12]. In this investigation, urine samples obtained in a long-term toxicity study, at different dose levels, were analyzed using reversed-phase LC-ESI/TOF MS analysis and negative ion detection. Using PCA it was possible to show clear differences between dose groups and the controls. Five unique *m/z* values were identified from the PCA loading plots as strong contributors to the observed clustering, although specific metabolites were not assigned. Because of their initial work, these researchers have introduced substantial refinement in their approach including the use of ultra high-pressure LC to increase the peak capacity [19]. They have also shown

the utility of MS to allow drugs and their metabolites to be identified [100]. Other examples of metabolic profiling using LC–ESI/MS have predominantly involved toxicology. Several studies focused on the effects of nephrotoxins on the urinary metabolite profiles of rats that had been with administered mercuric chloride [21], cyclosporine [22], gentamicin [23], D-serine [101]. In another case, Lafaye and co-workers employed LC–ESI/MS to investigate heavy metal exposure in rodents over a 3-month period [102]. Differences were noted between the two heavy metal toxins in terms of the metabolites seen with dosing. The authors utilized the MS/MS capability of this instrument to derive structural information for specific metabolites. Then they reported a LC/ESI-MS method for profiling a wide range of structurally different sulfoconjugated compounds in urine and its application in the characterization of biomarkers for heavy metal toxicity in rat urine [103]. Comparative analysis of the chromatographic fingerprints of urine from control and uranium- and cadmium-treated rats showed several variations in the chromatographic pattern of the sulfoconjugates. Diagnostic *m/z* ratios were confirmed by analyzing individual urine samples and one of the observed variations seemed to be specific to uranium toxicity. Using a different approach Gamache et al. [104] provided structural insight for "redox active" metabolites in rat urine by incorporating a parallel electrochemical array on-line with LC–ESI-TOF/MS to examine the influence of known kidney and liver toxicants. This technology may indeed be useful in toxicological investigations since many routes of toxicity involve redox active species. Idborg-Bjorkman et al. reported a quadrupole-LC–ESI/MS method for the investigation of drug-induced phospholipidosis by the antidepressant citalopram [105]. In this study, urine samples from rats exposed to citalopram were analyzed using solid-phase extraction and HPLC–MS analysis detecting negative ions. A number of potential biomarkers of phospholipidosis in rats are discussed.

4.2. Diseases biomarker discovery

Many examples exist in the literature on the use of NMR-based metabolic profiling to aid human disease diagnosis, such as the use of plasma to study diabetes, cerebrospinal fluid for investigating Alzheimer's disease, synovial fluid for osteoarthritis, seminal fluid for male infertility, and urine in the investigation of drug overdose, renal transplantation, and various renal diseases [106]. The investigations of LC–MS-based metabonomcis method to disease diagnosis are still in the early stage. Authors' group developed a LC–MS combination with multivariate statistical analysis method to investigate the plasma phospholipids metabolic profiling of type 2 diabetes mellitus (DM2) [107]. PCA and PLS-DA models were tested and compared in class separation between the DM2 and control. Using this method, the DM2 patients could be differentiated from the control and the potential biomarkers were identified. The plasma phospholipid metabolic profile in mouse IgA nephropathy was also investigated using the similar LC/MS method by our group [108]. It was found that expression of intercellular adhesion molecule-1 (ICAM-1) in the glomeruli had a significant correlation with proteinuria in mouse IgA nephropathy. The association

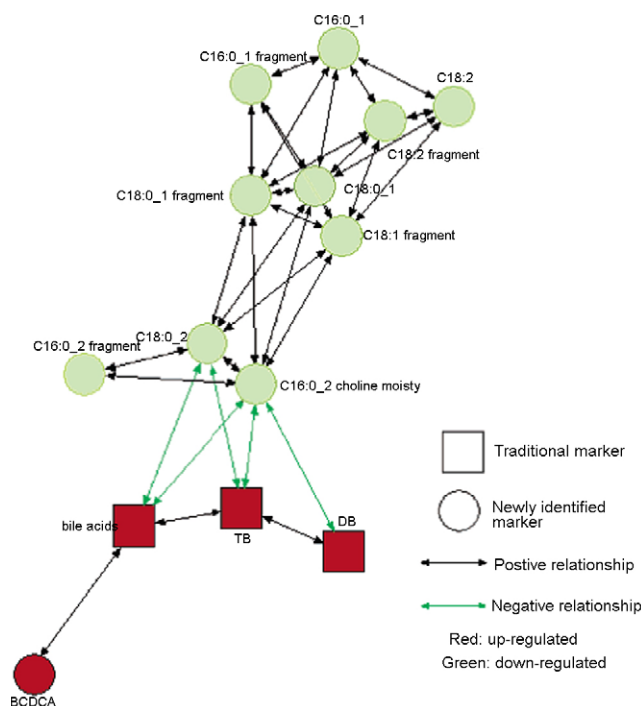

Fig. 7. Correlation network of biomarkers with $|C_{ij}| > 0.8$. The high relationship between the parent ions such as C16:0 and the corresponding fragment ions further indicated the reliability of the identification. Reproduced with permission from Yang et al. [109].

between plasma phospholipids and expression of ICAM-1 in the glomeruli of IgA nephropathy suggested C18:0/CI8:0 PS (phosphatidylserine), C18:0/C22:5 PS (phosphatidylserine) and C18:0/C20:4 PI (phosphatidylinositol) were possible biomarkers of IgA nephropathy. We also developed a HPLC–MS-based metabonomis technique and applied it to the investigation of acute deterioration of liver function in chronic hepatitis B to find the potential biomarkers [109]. Several potential biomarkers were identified: lysophosphatidyl choline (LPC) C18:0, LPC C16:0, LPC C18:1, LPC C18:2, and glycochenodeoxycholic acid (GCDCA) (or its isomer glycodeoxycholic acid (GDCA)). Combined the marker list with the traditional marker data, the correlation analysis was performed, it is found from Fig. 7 that GCDCA has a very good relationship with bile acids, total bilirubin (TB) and direct bilirubin (DB), their concentrations were up-regulated. In reverse, the concentrations of LPCs in chronic hepatitis B patients hospitalized for acute deterioration of liver function were down-regulated. The fragments and their corresponding parent ions were also highly correlated, which is helpful for the structure identification.

5. Future prospects

It is out of question that HPLC–MS technique will continue as an important technology for the identification and the quantification of biomarkers. The application of HPLC–MS-based metabonomics analysis has inevitably given the great benefit to pharmaceutical industry. The studies reported in the literatures have demonstrated considerable promise for HPLC–MS-based analytical strategies for metabonomic and biomarker research.

The future technological, methodological, and informatics advances will increase its impact on metabonomics investigations and biomarker discovery. The improvement in the mass accuracy and resolution of mass spectrometers will enable better prediction of metabolite biomarker identities. Many TOF and FT-MS mass spectrometers already provide 3 ppm mass accuracy and further improvement can be expected. The sensitivity of mass spectrometers will continue to increase and further broaden its applicability to trace analyte detection in complex biological media. With the improvement of mass spectrometry technologies, methodologies and informatics will be developed to facilitate new capabilities. Increased use of immunoaffinity [110], multidimensional chromatography, and other novel separations methods coupled with mass spectrometry will further increase the sensitivity and dynamic range of trace biomarker detection. The use of capillary LC and UPLC methods will improve the coverage of the metabolome provided by HPLC–MS as a result of reduced ion suppression. The difficulty in exchanging a spectra library from LC–MS analysis is a major drawback of that technology, particularly for metabolite profiling, where the identification of a large number of compounds is desirable. Increasing our knowledge in ionization mechanisms in LC–MS may lead to the possibility to exchange spectra libraries between laboratories and research groups and that exchange will have a substantial impact on the entire field of metabolome analysis. Increased incorporation of stable isotope labeling into biomarker identification and quantification methods will greatly improve assay validation and biomarker qualification. Software and database advancements will complement advances in mass spectrometry technology and methods. A metabolite database that allows quick mass spectrometric identification of endogenous metabolites may advance metabonomics to the user level of genomics, and bioinformatics advances that allow systems biology understanding of RNA, protein, and metabolite changes in the same biological systems will lead to faster evaluation of novel biomarker candidates.

Acknowledgments

The work has been supported by the foundation (No. 20425516) for Distinguished Young Scholars and the foundation (No. 20675082) from National Natural Science Foundation of China, Sino-German Center for Research Promotion (DFG and NSFC, GZ 364), National Basic Research Program (2006CB503902) of State Ministry of Science and Technology of China and the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX1-YW-02, K2006A13).

References

[1] J.K. Nicholson, J.C. Lindon, E. Holmes, Xenobiotica 29 (1999) 1181.
[2] J.C. Lindon, E. Holmes, M.E. Bollard, E.G. Stanley, J.K. Nicholson, Biomarkers 9 (2004) 1.
[3] J.K. Nicholson, J. Connelly, J.C. Lindon, E. Holmes, Nat. Rev. Drug Discov. 1 (2002) 153.
[4] M.L. Anthony, B.C. Sweatman, C.R. Beddell, J.C. Lindon, J.K. Nicholson, Mol. Pharmacol. 46 (1994) 199.
[5] M.L. Anthony, V.S. Rose, J.K. Nicholson, J.C. Lindon, J. Pharm. Biomed. Anal. 13 (1995) 205.
[6] D.G. Robertson, M.D. Reily, R.E. Sigler, D.F. Wells, D.A. Paterson, T.K. Braden, Toxicol. Sci. 57 (2000) 326.
[7] O. Fiehn, Plant Mol. Biol. 48 (2002) 155.
[8] O. Fiehn, J. Kopka, P. Dormann, T. Altmann, R.N. Trethewey, L. Willmitzer, Nat. Biotechnol. 18 (2000) 1157.
[9] J.C. Lindon, E. Holmes, J.K. Nicholson, Anal. Chem. 75 (2003) 384A.
[10] J.T. Brindle, H. Antti, E. Holmes, G. Tranter, J.K. Nicholson, H.W.L. Bethell, S. Clarke, P.M. Schofield, E. McKilligin, D.E. Mosedale, D.J. Grainger, Nat. Med. 8 (2002) 1439.
[11] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass Spectrom. Rev. 26 (2007) 51.
[12] R.S. Plumb, C.L. Stumpf, M.V. Gorenstein, J.M. Castro-Perez, G.J. Dear, M. Anthony, B.C. Sweatman, S.C. Connor, J.N. Haselden, Rapid Commun. Mass Spectrom. 16 (2002) 1991.
[13] I.D. Wilson, J.K. Nicholson, J. Castro-Perez, J.H. Granger, K.A. Johnson, B.W. Smith, R.S. Plumb, J. Proteome Res. 4 (2005) 591.
[14] I.D. Wilson, R. Plumb, J. Granger, H. Major, R. Williams, E.A. Lenz, J. Chromatogr. B 817 (2005) 67.
[15] J.L. Griffin, Curr. Opin. Chem. Biol. 7 (2003) 648.
[16] E.M. Lenz, J. Bright, I.D. Wilson, S.R. Morgan, A.F.P. Nash, J. Pharm. Biomed. Anal. 33 (2003) 1103.
[17] I. Pelczer, Curr. Opin. Drug Discov. Dev. 8 (2005) 127.
[18] W.B. Dunn, D.I. Ellis, Trac-Trends Anal. Chem. 24 (2005) 285.
[19] R.S. Plumb, J.H. Granger, C.L. Stumpf, K.A. Johnson, B.W. Smith, S. Gaulitz, I.D. Wilson, J. Castro-Perez, Analyst 130 (2005) 844.
[20] S. Wagner, K. Scholz, M. Donegan, L. Burton, J. Wingate, W. Volkel, Anal. Chem. 78 (2006) 1296.
[21] E.M. Lenz, J. Bright, R. Knight, I.D. Wilson, H. Major, Analyst 129 (2004) 535.
[22] E.M. Lenz, J. Bright, R. Knight, I.D. Wilson, H. Major, J. Pharm. Biomed. Anal. 35 (2004) 599.
[23] E.M. Lenz, J. Bright, R. Knight, F.R. Westwood, D. Davies, H. Major, I.D. Wilson, Biomarkers 10 (2005) 173.
[24] E.M. Lenz, I.D. Wilson, J. Proteome Res. 6 (2007) 443.
[25] H. Idborg, L. Zamani, P.O. Edlund, I. Schuppe-Koistinen, S.P. Jacobsson, J. Chromatogr. B 828 (2005) 9.
[26] A.D. Delinsky, D.C. Delinsky, S. Muralidhara, J.W. Fisher, J.V. Bruckner, M.G. Bartlett, Rapid Commun. Mass Spectrom. 19 (2005) 1075.
[27] R. Pisano, M. Breda, S. Grassi, C.A. James, J. Pharm. Biomed. Anal. 38 (2005) 738.
[28] V.V. Tolstikov, A. Lommen, K. Nakanishi, N. Tanaka, O. Fiehn, Anal. Chem. 75 (2003) 6737.
[29] J. Granger, R. Plumb, J. Castro-Perez, I.D. Wilson, Chromatographia 61 (2005) 375.
[30] R.S. Plumb, K.A. Johnson, P. Rainville, J.P. Shockcor, R. Williams, J.H. Granger, I.D. Wilson, Rapid Commun. Mass Spectrom. 20 (2006) 2800.
[31] R.S. Plumb, K.A. Johnson, P. Rainville, B.W. Smith, I.D. Wilson, J.M. Castro-Perez, J.K. Nicholson, Rapid Commun. Mass Spectrom. 20 (2006) 2234.
[32] P.Y. Yin, X.J. Zhao, Q.R. Li, J.S. Wang, J.S. Li, G.W. Xu, J. Proteome Res. 5 (2006) 2135.
[33] X.J. Zhao, W.Z. Wang, J.S. Wang, J. Yang, G.W. Xu, J. Sep. Sci. 29 (2006) 2444.
[34] K.S. Boos, A. Rudolphi, S. Vielhauer, A. Walfort, D. Lubda, F. Eisenbeiss, Fresenius J. Anal. Chem. 352 (1995) 684.
[35] A. Rudolphi, S. Vielhauer, K.S. Boos, D. Seidel, I.M. Bathge, H. Berger, J. Pharm. Biomed. Anal. 13 (1995) 615.
[36] K.S. Boos, J. Lintelmann, A. Kettrup, J. Chromatogr. 600 (1992) 189.
[37] Y.F. Zheng, G.W. Xu, J. Yang, X.J. Zhao, T. Pang, H.W. Kong, J. Chromatogr. B 819 (2005) 85.
[38] G.W. Xu, Y. Wang, China Pat., Appl. No. 200610134027.5. (2006).
[39] G.W. Xu, Y. Wang, China Pat., Appl. No. 200610134028.X. (2006).
[40] Y. Wang, X. Lu, G.W. Xu, Proceedings of Dalian International Symposia and Exhibition on Chromatography, Chemical Industry Press, Dalian, China, 2007, p. 213.
[41] S.J. Gaskell, J. Mass Spectrom. 32 (1997) 677.

[42] D.V. Huhman, L.W. Sumner, Phytochemistry 59 (2002) 347.
[43] C.G. Enke, Anal. Chem. 69 (1997) 4885.
[44] R. Goodacre, S. Vaidyanathan, G. Bianchi, D.B. Kell, Analyst 127 (2002) 1457.
[45] K.O. Boernsen, S. Gatzek, G. Imbert, Anal. Chem. 77 (2005) 7255.
[46] L.A. Leuthold, C. Grivet, M. Allen, M. Baumert, G. Hopfgartner, Rapid Commun. Mass Spectrom. 18 (2004) 1995.
[47] J.T. Kapron, E. Pace, C.K. Van Pelt, J. Henion, Rapid Commun. Mass Spectrom. 17 (2003) 2019.
[48] R.F. Staack, E. Varesio, G. Hopfgartner, Rapid Commun. Mass Spectrom. 19 (2005) 618.
[49] G.F. Wang, Y. Hsieh, W.A. Korfmacher, Anal. Chem. 77 (2005) 541.
[50] A. Raffaelli, A. Saba, Mass Spectrom. Rev. 22 (2003) 318.
[51] J.E. Dally, J. Gorniak, R. Bowie, C.M. Bentzley, Anal. Chem. 75 (2003) 5046.
[52] H.R. Morris, T. Paxton, A. Dell, J. Langhorne, M. Berg, R.S. Bordoli, J. Hoyes, R.H. Bateman, Rapid Commun. Mass Spectrom. 10 (1996) 889.
[53] G. Hopfgartner, C. Husser, M. Zell, J. Mass Spectrom. 38 (2003) 138.
[54] J.C.Y. Le Blanc, J.W. Hager, A.M.P. Ilisiu, C. Hunter, F. Zhong, I. Chu, Proteomics 3 (2003) 859.
[55] J.E.P. Syka, J.A. Marto, D.L. Bai, S. Horning, M.W. Senko, J.C. Schwartz, B. Ueberheide, B. Garcia, S. Busby, T. Muratore, J. Shabanowitz, D.F. Hunt, J. Proteome Res. 3 (2004) 621.
[56] G.J. Dear, J. Ayrton, R. Plumb, I.J. Fraser, Rapid Commun. Mass Spectrom. 13 (1999) 456.
[57] Y.Q. Xia, J.D. Miller, R. Bakhtiar, R.B. Franklin, D.Q. Liu, Rapid Commun. Mass Spectrom. 17 (2003) 1137.
[58] A. Makarov, E. Denisov, O. Lange, S. Horning, J. Am. Soc. Mass Spectrom. 17 (2006) 977.
[59] A. Makarov, E. Denisov, A. Kholomeev, W. Baischun, O. Lange, K. Strupat, S. Horning, Anal. Chem. 78 (2006) 2113.
[60] A.G. Marshall, C.L. Hendrickson, G.S. Jackson, Mass Spectrom. Rev. 17 (1998) 1.
[61] I.J. Amster, J. Mass Spectrom. 31 (1996) 1325.
[62] U. Bahr, M. Karas, F. Hillenkamp, Fresenius J. Anal. Chem. 348 (1994) 783.
[63] K. Strupat, M. Karas, F. Hillenkamp, C. Eckerskorn, F. Lottspeich, Anal. Chem. 66 (1994) 464.
[64] E. Nordhoff, A. Ingendoh, R. Cramer, A. Overberg, B. Stahl, M. Karas, F. Hillenkamp, P.F. Crain, Rapid Commun. Mass Spectrom. 6 (1992) 771.
[65] G. Sun, K. Yang, Z.D. Zhao, S.P. Guan, X.L. Han, R.W. Gross, Anal. Chem. 79 (2007) 6629.
[66] P. Jonsson, S.J. Bruce, T. Moritz, J. Trygg, M. Sjostrom, R. Plumb, J. Granger, E. Maibaum, J.K. Nicholson, E. Holmes, H. Antti, Analyst 130 (2005) 701.
[67] C.D. Broeckling, I.R. Reddy, A.L. Duran, X.C. Zhao, L.W. Sumner, Anal. Chem. 78 (2006) 4334.
[68] M. Katajamaa, J. Miettinen, M. Oresic, Bioinformatics 22 (2006) 634.
[69] M. Katajamaa, M. Oresic, BMC Bioinformatics 6 (2005) 12.
[70] C.A. Smith, E.J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, Anal. Chem. 78 (2006) 779.
[71] J. Yang, G.W. Xu, W.F. Zheng, H.W. Kong, C. Wang, X.J. Zhao, T. Pang, J. Chromatogr. A 1084 (2005) 214.
[72] H. Idborg, L. Zamani, P.O. Edlund, I. Schuppe-Koistinen, S.P. Jacobsson, J. Chromatogr. B 828 (2005) 14.
[73] G. Schlotterbeck, A. Ross, F. Dieterle, H. Senn, Pharmacogenomics 7 (2006) 1055.
[74] H. Martens, T. Naes (Eds.), Multivariate Calibration, John Wiley & Sons Inc., New York, 1989, p. 93.
[75] J. Trygg, S. Wold, J. Chemom. 16 (2002) 119.
[76] R. Goodacre, J. Exp. Bot. 56 (2005) 245.
[77] J. Trygg, E. Holmes, T. Lundstedt, J. Proteome Res. 6 (2007) 469.
[78] http://www.nist.gov/srd/nist1.htm.
[79] http://www.genome.jp/kegg/ligand.html.
[80] http://biocyc.org.
[81] http://www.metabolome.jp.
[82] http://www.ualberta.ca/_gjones/mslib.htm.
[83] http://metlin.scripps.edu/.
[84] C.A. Smith, G. O'Maille, E.J. Want, C. Qin, S.A. Trauger, T.R. Brandon, D.E. Custodio, R. Abagyan, G. Siuzdak, Ther. Drug Monit. 27 (2005) 747.
[85] http://www.metabolic-database.com.
[86] http://www.lipidmaps.org/data/index.html.
[87] http://sphingomap.org/.
[88] http://lipidbank.jp/.
[89] http://www.hmdb.ca/.
[90] D.S. Wishart, D. Tzur, C. Knox, R. Eisner, A.C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D.D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G.E. Duggan, G.D. MacInnis, A.M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B.D. Sykes, H.J. Vogel, L. Querengesser, Nucleic Acids Res. 35 (2007) D521.
[91] C. Prakash, C.L. Shaffer, A. Nedderman, Mass Spectrom. Rev. 26 (2007) 340.
[92] S.G. Ma, S.K. Chowdhury, K.B. Alton, Curr. Drug Metab. 7 (2006) 503.
[93] A.E.F. Nassar, R.E. Talaat, Drug Discov. Today 9 (2004) 317.
[94] R.E. Williams, E.M. Lenz, J.S. Lowden, M. Rantalainen, I.D. Wilson, Mol. Biosyst. 1 (2005) 166.
[95] A.W.T. Bristow, K.S. Webb, A.T. Lubben, J. Halket, Rapid Commun. Mass Spectrom. 18 (2004) 1447.
[96] R. Jansen, G. Lachatre, P. Marquet, Clin. Biochem. 38 (2005) 362.
[97] A. Pelander, I. Ojanpera, S. Laks, I. Rasanen, E. Vuori, Anal. Chem. 75 (2003) 5710.
[98] J.M. Halket, D. Waterman, A.M. Przyborowska, R.K.P. Patel, P.D. Fraser, P.M. Bramley, J. Exp. Bot. 56 (2005) 219.
[99] D.G. Robertson, Toxicol. Sci. 85 (2005) 809.
[100] R.S. Plumb, C.L. Stumpf, J.H. Granger, J. Castro-Perez, J.N. Haselden, G.J. Dear, Rapid Commun. Mass Spectrom. 17 (2003) 2632.
[101] R.E. Williams, H. Major, E.A. Lock, E.M. Lenz, I.D. Wilson, Toxicology 207 (2005) 179.
[102] A. Lafaye, C. Junot, B. Ramounet-Le Gall, P. Fritsch, J.C. Tabet, E. Ezan, Rapid Commun. Mass Spectrom. 17 (2003) 2541.
[103] A. Lafaye, C. Junot, B. Ramounet-Le Gall, P. Fritsch, E. Ezan, J.C. Tabet, J. Mass Spectrom. 39 (2004) 655.
[104] P.H. Gamache, D.F. Meyer, M.C. Granger, I.N. Acworth, J. Am. Soc. Mass Spectrom. 15 (2004) 1717.
[105] H. Idborg-Bjorkman, P.O. Edlund, O.M. Kvalheim, I. Schuppe-Koistinen, S.P. Jacobsson, Anal. Chem. 75 (2003) 4784.
[106] J.C. Lindon, E. Holmes, J.K. Nicholson, Pharm. Res. 23 (2006) 1075.
[107] C. Wang, H.W. Kong, Y.F. Guan, J. Yang, J.R. Gu, S.L. Yang, G.W. Xu, Anal. Chem. 77 (2005) 4108.
[108] L.W. Jia, C. Wang, H.W. Kong, Z.W. Cai, G.W. Xu, Metabolomics 2 (2006) 95.
[109] J. Yang, X.J. Zhao, X.L. Liu, C. Wang, P. Gao, J.S. Wang, L.J. Li, J.R. Gu, S.L. Yang, G.W. Xu, J. Proteome Res. 5 (2006) 554.
[110] J.W. Sen, H.R. Bergen, N.H.H. Heegaard, Anal. Chem. 75 (2003) 1196.