DOI: 10.1002/cbic.200500151

The Expanding Role of Mass Spectrometry in Metabolite Profiling and Characterization

Elizabeth J. Want, Benjamin F. Cravatt, and Gary Siuzdak*[a]

Mass spectrometry has a strong history in drug-metabolite analysis and has recently emerged as the foremost technology in endogenous metabolite research. The advantages of mass spectrometry include a wide dynamic range, the ability to observe a diverse number of molecular species, and reproducible quantitative analysis. These attributes are important in addressing the issue of metabolite profiling, as the dynamic range easily exceeds nine orders of magnitude in biofluids, and the diversity of species ranges from simple amino acids to lipids to complex carbohydrates. The goals of the application of mass spectrometry range from basic biochemistry to clinical biomarker discovery with challenges in generating a comprehensive profile, data analysis, and structurally characterizing physiologically important metabolites. The precedent for this work has already been set in neonatal screening, as blood samples from millions of neonates are tested routinely by mass spectrometry as a diagnostic tool for inborn errors of metabolism. In this review, we will discuss the background from which contemporary metabolite research emerged, the techniques involved in this exciting area, and the current and future applications of this field.

Introduction

The power of mass spectrometry in endogenous metabolite research derives from its proven success in drug-metabolite analysis and pharmacokinetic studies.[1–9] Mass spectrometry also facilitates the identification of previously uncharacterized metabolites.[10,11] Mass spectrometry is already well established as a quantitative tool for small molecules, unlike its application to proteomics. This ability, and the fact that the metabolic profile is a reflection of enzymatic activity, allow for the direct monitoring of both normal biochemical events and perturbations that lead to disease. For example, when a defective enzyme is produced with impaired or even absent activity, it can result in a change in the abundance of a particular metabolite. Currently ~200 inherited enzymatic disorders can be characterized by perturbations in their associated metabolic profiles,[12–22] and, clinically, over 30 disorders of the metabolism of organic acids, amino acids, fatty acids, and steroid hormones are monitored worldwide by using mass spectrometry.[16] Additionally, enzymatic activity can be influenced by external forces, such as drugs or infection, resulting in changes in metabolic profiles (Figure 1). The largest focus of metabolite studies has been on biofluids such as serum, plasma, and urine, as these provide readily accessible, noninvasive sources to examine the metabolic response(s) of the organism.

Recent interest in metabolite profiling is derived from the ability to perform more comprehensive metabolite analyses with new liquid chromatography/mass spectrometry (LC/MS) technology, the need to understand the biochemical events of cells and tissues, and the potential for identifying both disease and pharmaceutical biomarkers.

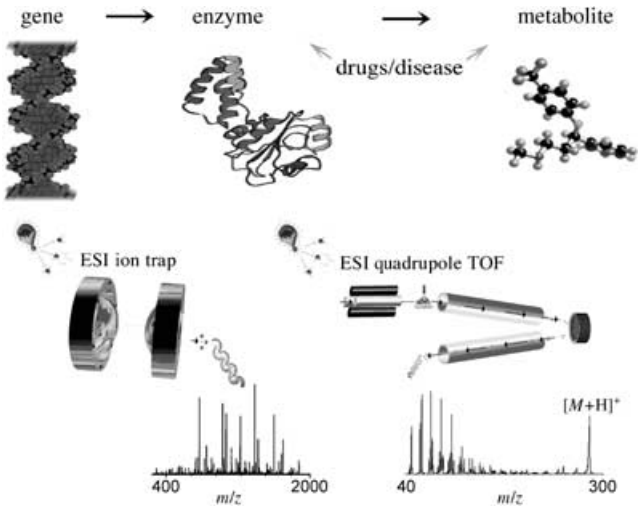

Figure 1. Simplified overview of the concepts involved in metabolite profiling. Alterations in metabolite levels reflect the activity of their corresponding enzymes. Additional factors that affect metabolite profiles include drug intake or the onset of a disease. Two of the common mass spectrometry approaches for proteomics (ESI ion trap) and metabolite profiling (ESI quadrupole TOF) are shown.

1. Instrumentation

In the 1970s and 1980s, gas chromatography mass spectrometry (GC/MS) was the most commonly used method for small-molecule analysis,[23] and is still used today for the detection of many metabolic disorders.[16] However, in the 1990s, GC/MS was largely superseded by high-field proton NMR, which, until

[a] *Dr. E. J. Want, Prof. B. F. Cravatt, Prof. G. Siuzdak*
Department of Molecular Biology and The Center for Mass Spectrometry
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, California 92037 (USA)
Fax: (+1) 858-784-9496
E-mail: siuzdak@scripps.edu

recently, remained the analytical method of choice.[24] Liquid chromatography combined with mass spectrometry (LC/MS) has demonstrated several advantages over NMR, including greater sensitivity and dynamic range. In this section, we will discuss the background from which contemporary metabolite research emerged and the current applications of this field. We will also discuss the bioinformatics tools and mathematical modeling techniques, such as principal-component and cluster analysis, used to maximize information recovery and to aid in the interpretation of the very large multivariate data sets typically obtained in metabolite-profiling studies.

NMR

NMR has been applied to metabolite-profiling studies[25–27] in areas as diverse as plant metabolism,[28,29] Duchenne muscular dystrophy,[30] bioavailability and metabolic responses of rats to epicatechin,[31] hypertension,[32] and acetaminophen toxicity.[33] The primary advantage of NMR is its ability to measure analytes in biofluids quickly and accurately, without the need for initial processing or separation. Over recent years, improvements have included higher resolution, lower instrument cost, and the addition of stop–flow chromatography on fractions of samples. However, a major weakness of NMR is that it has a poor dynamic range (100–1000) that results in only the most abundant components being observed (Table 1).

GC/MS

Historically, the combination of high-resolution capillary gas chromatography with mass spectrometry (GC/MS) was the most widely used application in metabolite research and disease diagnosis. GC/MS enabled the identification of key small molecules—such as fatty acids, amino acids, and organic acids—in biofluids, particularly in urine and blood.[23,34,35] GC/MS has been instrumental in providing diagnostic information for many inherited diseases, including numerous metabolic disorders. These include disorders of the metabolism of amino acids,[36–39] thyroid hormones,[40–42] bile acids,[43,44] steroids,[45–47] organic acids,[48] and fatty acids,[49,50] with 20–30 disorders of the latter alone having been characterized.[16]

An important example of quantitative mass spectrometry metabolite analysis is the measurement of phenylalanine and tyrosine in the diagnosis of phenylketonuria (PKU).[16] Here, GC and GC/MS were used to establish the normal metabolite pattern and ultimately identify the disease state (Figure 2). Many clinical laboratories around the world still use GC/MS for the diagnosis of some of these metabolic diseases, particularly those disorders of organic acids, in which, at present, LC/MS techniques have offered little advancement.[16] This is due in part to the lower resolution of LC and the fact that electrospray ionization (ESI) of organic acids produces negative ions that are not detected with the same degree of sensitivity as positive ions. Derivatization of these molecules would aid in their analysis, but offers no improvement over the well-established GC/MS techniques. Therefore, diseases such as organic acidemias[48] and many steroid-metabolism disorders[45–47] are still detected by using GC/MS.[16] Shackleton and colleagues pioneered the technique of urinary steroid profiling using GC and GC/MS in the late 1960s and 1970s.[45,51] This procedure is now in place in many laboratories worldwide and is invaluable in the detection of inborn errors of steroid metabolism, such as congenital adrenal hyperplasia,[46] and has shown potential in screening for Smith–Lemli–Opitz syndrome, in which a fundamental defect in cholesterol metabolism exists.[52]

GC/MS has also been applied successfully to metabolite profiling for plant functional genomics,[53–57] which, combined with principal-component analysis (PCA), has been used to significantly improve upon existing approaches and understanding.

However, there are several disadvantages to using GC/MS that give it limited applicability to metabolite profiling. These include convoluted sample preparation that involves metabolite extraction as well as derivatization to improve volatility, lengthy analysis time, and the limits on the size and type of molecule that can be analyzed (nonvolatile, polar macromolecules are unsuitable). LC/MS with an electrospray interface has become a more popular choice for these analyses and studies

Table 1. Advantages and disadvantages of using GC/MS, NMR, ESI-LC/MS and flow injection analysis (FIA) with ESI in metabolite-profiling studies.

	LC/MS	FIA	GC/MS	NMR
sample preparation	extraction	extraction	extraction and chemical modification	typically none
chromatographic separation	medium-resolution separation	no separation	high-resolution separation	no separation
sensitivity	millimolar to nanomolar	millimolar to micromolar	millimolar to nanomolar	millimolar to high micromolar
dynamic range	10^6	10^4	10^6	10^3
speed	slow (5 to 90 min)	rapid (1 to 5 min)	slow (~30 min)	rapid (1 to 5 min)
quantitative accuracy	±10%	±10%	±10%	±10%
significant advantages	soft ionization; large mass range	data in one spectrum; fast	high resolution; EI-MS library available	no sample preparation
significant disadvantages	speed of analyses	signal suppression from multiple components	significant sample preparation with chemical modification, slow analysis time, harsh ionization, and limited number of molecules can be analyzed	poor sensitivity and dynamic range; some chemical classes not detected

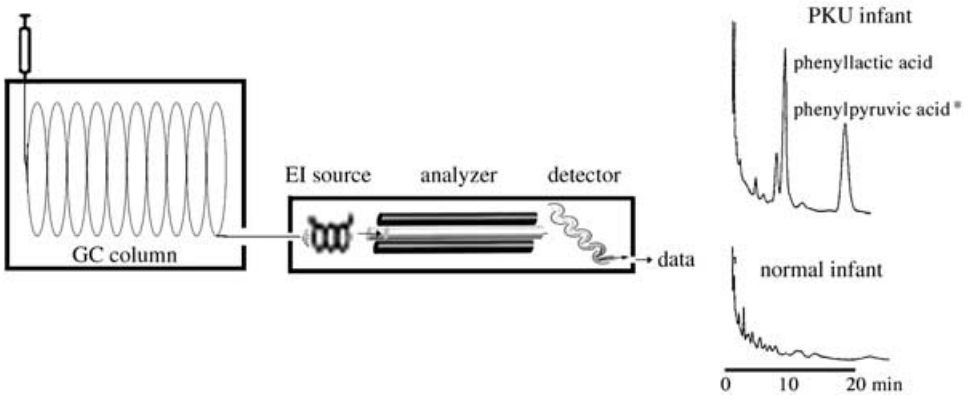

Figure 2. GC and GC/MS for PKU diagnosis. GC ion chromatograms of urinary extracts from a PKU infant and a healthy infant, showing phenyllactic acid and phenylpyruvic acid in the PKU infant.

for new biomarkers.[35;58] This technique is advantageous over GC/MS in that sample preparation and analysis are relatively simple, providing access to metabolites of more diverse chemical structure and size.

LC/MS

Liquid chromatography electrospray ionization mass spectrometry (LC/ESI-MS; Figure 3) is now the most common approach toward metabolite-profiling studies. ESI offers many advantages over other ionization techniques, including the ability to analyze low- and high-mass compounds, excellent quantitative capabilities and reproducibility, high sensitivity, simple sample preparation, amenability to automation, soft ionization, and the absence of matrix (as required for matrix-assisted laser desorption/ionization mass spectrometry (MALDI)).[59]

In its simplest form, ESI can be quite effective even without separation, especially when combined with tandem mass spectrometry (MS/MS). Direct ESI-MS/MS applications to metabolite screening emerged in the early 1990s, and today over thirty five diseases, including the screening of newborns for phenylketonuria, are monitored directly by using ESI-MS/MS (Figure 4).[12,60] In fact, due to the importance of initial GC/MS results and the ESI-MS/MS clinical assay, Neo Gen Screening was founded in 1994 (currently Pediatrix Screening http://www.pediatrixscreening.com). Worldwide, blood samples from over two million newborns have been ESI screened for at least 30 amino acid, organic acid, and fatty acid oxidation disorders (Table 2).[12–21] Among these diseases, PKU is a useful example for demonstrating the utility of ESI-MS/MS (Figure 4).

The utility of ESI lies in its ability to generate gas-phase ions directly from the liquid phase; this establishes the technique as a convenient mass-analysis platform for both liquid chromatography and automated sample analysis. Separation of the thousands of molecules present in biofluids can reduce ESI ion suppression[61–63] by decreasing the number of competing analytes entering the mass spectrometer ion source at any one time. These factors result in a selective approach that allows for both quantitation and structural information, while sensitivities in the pg/mL range can be achieved readily.[64] Given these advantages, LC/MS techniques have replaced some of the traditional specialized clinical laboratory methods[65,66] that relied on biological, immunological, and fluorometric techniques.[23]

A challenge in metabolite profiling is that potential biomarkers might be present in the biofluid in low abundance, thus requiring sensitive analysis techniques in order to detect them. One approach is nanoESI liquid chromatography, which has already proved important in proteomics studies.[67,68]

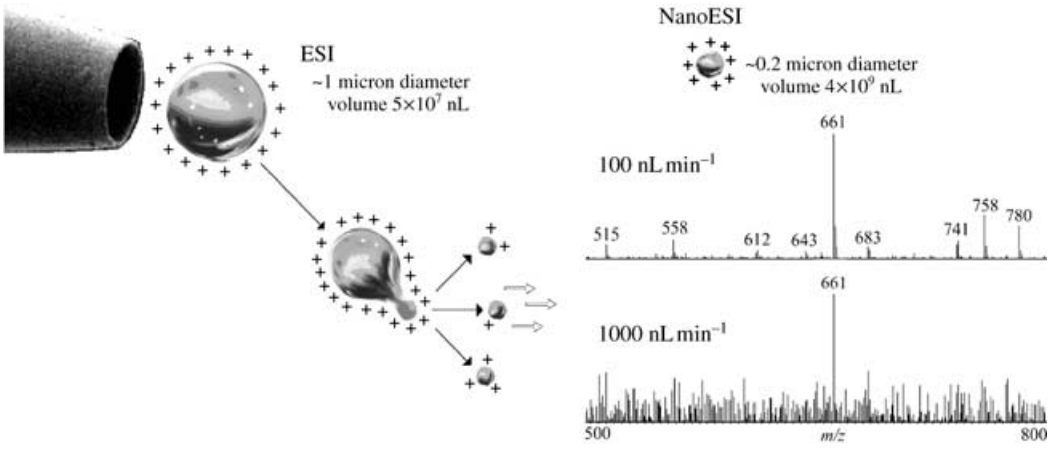

Figure 3. The small droplet size in nanoESI leads to less signal suppression and greater coverage of a solution's metabolites. LC/MS spectra of methanol-extracted serum samples at flow rates of 100 and 1000 nL min^{-1} show a significant improvement in the signal-to-noise ratio at the lower flow rate, thus resulting in enhanced detection of metabolites.

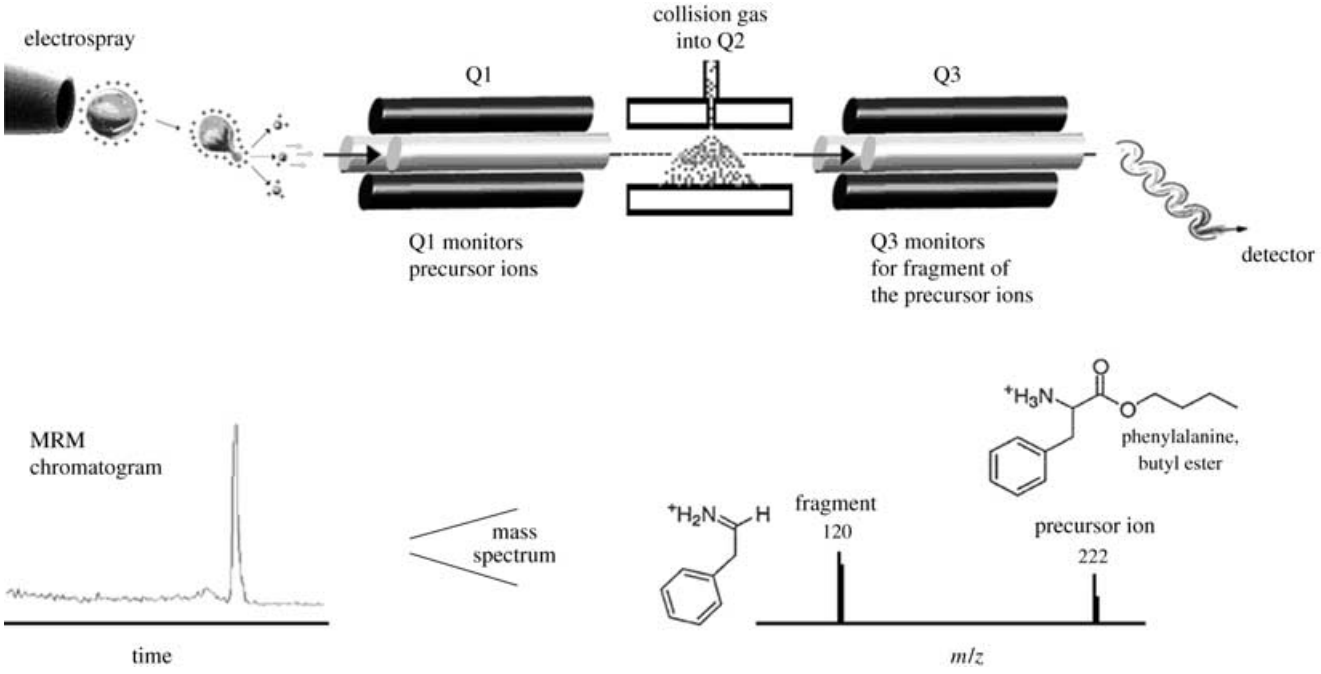

Figure 4. Multiple-reaction monitoring (MRM) with a triple quadrupole. LC MS/MS experiments are now standard for monitoring over 35 different diseases in neonates. Here, an example of monitoring for phenlyketonuria (PKU) with a dried blood spot from an infant by using a triple quadrupole mass analyzer is shown.

Table 2. The major metabolic disorders screened for in newborns using mass spectrometry.[a]

Fatty acid oxidation	Organic acid	Amino acid
carnitine/acylcarnitine translocase deficiency (translocase)	3-hydroxy-3-methylglutaryl-CoA lyase deficiency (HMG)	argininemia
carnitine palmitoyl transferase deficiency type I (CPT-I)	glutaric acidemia type I (GA I)	argininosuccinic aciduria (ASA lyase deficiency)
3-hydroxy long chain acyl-CoA dehydrogenase deficiency (LCHAD)	isobutyryl-CoA dehydrogenase deficiency	· acute onset
2,4-dienoyl-CoA reductase deficiency	isovaleric acidemia (IVA)	· late onset
medium chain acyl-CoA dehydrogenase deficiency (MCAD)	· acute onset	carbamoylphosphate synthetase deficiency (CPS Def.)
multiple acyl-CoA dehydrogenase deficiency (MADD or glutaric acidemia-type II)	· chronic	citrullinemia (ASA synthetase deficiency)
neonatal carnitine palmitoyl transferase deficiency type II (CPT-II)	2-methylbutryl-CoA dehydrogenase deficiency	· acute onset
short-chain acyl-CoA dehydrogenase deficiency (SCAD)	3-methylcrotonyl-CoA carboxylase deficiency (3MCC deficiency)	· late onset
short-chain hydroxy acyl-CoA dehydrogenase deficiency (SCHAD)	3-methylglutaconyl-CoA hydratase deficiency	homocystinuria
trifunctional protein deficiency (TFP deficiency)	methylmalonic acidemias	hypermethioninemia
very long chain acyl-CoA dehydrogenase deficiency (VLCAD)	· methylmalonyl-CoA mutase deficiency 0	hyperammonemia, hyperornithinemia, homocitrullinemia syndrome (HHH)
	· methylmalonyl-CoA mutase deficiency +	hyperornithinemia with gyral atrophy
	· some adenosylcobalamin synthesis defects	maple syrup urine disease (MSUD)
	· maternal vitamin B12 deficiency	· classical MSUD
	mitochondrial acetoacetyl-CoA thiolase deficiency	· intermediate MSUD
	(3-ketothiolase deficiency.)	5-oxoprolinuria (pyroglutamic aciduria)
	propionic acidemia (PA)	phenylketonuria (PKU)
	· acute onset	· classical PKU
	· late onset	· hyperphenylalaninemia
	multiple-CoA carboxylase deficiency	· biopterin cofactor deficiencies (4)
	malonic aciduria	tyrosinemia
		transient neonatal tyrosinemia
		· tyrosinemia type I (Tyr I)
		· tyrosinemia type II (Tyr II)
		· tyrosimenia type III (Tyr III)

[a] List adapted from http://www.pediatrixscreening.com, with permission.

NanoLC/nano-ESI-MS is performed at much lower flow rates (~200 nL min^{-1}) than LC/ESI-MS (~300 μL min^{-1}; Figure 3); this improves the sensitivity and dynamic range, which is advantageous in metabolite analysis.[69–71] At the lower flow rates, the droplet size is significantly smaller; this means that as the droplet evaporates it reaches a high charge density more quickly than normal LC/ESI-MS. This is important because ESI is an evaporative process—charge density increases by vaporization—for larger droplets, the impurity concentration increases significantly before ionization occurs, ultimately inhibiting vaporization and concomitantly inhibiting ionization. In nanoLC/nano-ESI-MS, ions are produced with less evaporation and contain less-concentrated impurities, therefore increasing sensitivity (Figure 3) and ultimately offering a greater dynamic range in metabolite discovery.

Although the separation of complex mixture components makes LC/MS especially useful in the initial stages of metabolite research, there is a reduction in the speed of data acquisition and analysis. However, once a biomarker is identified, extraction combined with flow injection analysis (FIA) can be used to quantify the compound.[16] Here, the sample of interest is introduced directly into the mass spectrometer, without prior separation. This method is rapid, with an analysis time of 2 min or less per sample and can include automated procedures for peak assignment, making this method more advantageous than NMR.[72] Also, with this technique, all the information can be obtained in one spectrum. The primary disadvantage is that significant signal suppression can occur when a complex mixture is introduced, thus resulting in a loss of information and sensitivity.

Atmospheric-pressure chemical ionization (APCI) mass spectrometry is not widely used in metabolite-profiling studies. However, it can be a valuable technique for the analysis of neutral molecules, such as lipids.[73,74] Analogously to the ESI interface, APCI causes little or no fragmentation of the analyte and so is suitable for volatile and thermally stable compounds. APCI analysis of more easily ionizable molecules, such as phospholipids, can produce molecular and fragment ions complementary to those produced by ESI MS with collision-induced dissociation (CID). In addition, APCI provides a higher dynamic range than ESI and it is considered robust, easy to operate, and relatively more tolerant of higher buffer concentrations. As APCI can accommodate flow rates of up to 2.0 mL min^{-1}, it is considered mass-sensitive, as opposed to ESI, which is concentration-sensitive. However, one of the drawbacks of APCI is that, as it is a mass-sensitive device, no sensitivity gains are realized with smaller columns or lower flow rates.

Mass analyzers

The mass analyzer is critical to the performance of any mass spectrometer. Among the most commonly used are the quadrupole, quadrupole ion trap, time-of-flight, time-of-flight reflectron, and Fourier transform ion cyclotron resonance (FTMS). Quadrupole mass analyzers are the most common mass analyzers in existence today. These instruments tolerate relatively high pressures, have the capability of analyzing up to an *m/z* of 4000, and are relatively low-cost instruments. In order to perform tandem mass analysis with a quadrupole instrument, it is necessary to place three quadrupoles in series. Each quadrupole has a separate function: the first quadrupole (Q1) is used to scan across a preset *m/z* range and select an ion of interest. The second quadrupole (Q2), also known as the collision cell, focuses and transmits the ions while introducing a collision gas (argon or helium) into the flight path of the selected ion. The third quadrupole (Q3) serves to analyze the fragment ions generated in the collision cell (Q2; Figure 4).[59]

Quadrupole ion traps are useful in that a single ion species can be isolated by ejecting all others from the trap, and the isolated ions can be fragmented subsequently by collisional activation, and the fragments can be detected. A key advantage of quadrupole ion traps is that multiple CID experiments can be performed quickly without having multiple analyzers. Other advantages are their compact size and their ability to trap and accumulate ions to provide a better ion signal. They have a mass range up to ~4000 *m/z*. Less commonly used is the linear trap, but this instrument has advantages over the 3D trap in that the larger analyzer volume lends itself to a greater dynamic range and improved quantitative analysis. However, disadvantages of the ion trap include 1) the inability to perform high-sensitivity triple quadrupole-type precursor-ion scanning and neutral loss scanning experiments; 2) the upper limit on the ratio between precursor *m/z* and the lowest trapped fragment ion being ~0.3 (also known as the "one-third rule"); 3) a limited dynamic range because, when too many ions are in the trap, space charge effects diminish the performance of the ion-trap analyzer.[59]

The linear time-of-flight (TOF) is the simplest mass analyzer, with a virtually unlimited mass range. The TOF reflectron, now widely used for ESI, combines time-of-flight technology with an electrostatic mirror; this offers higher resolution (typically above 5000) than a simple TOF instrument. It has gained wide use due to its fast scanning capabilities (milliseconds), good mass range (up to *m/z* ~10000), and an accuracy in the order of 5 ppm. Quadrupole-TOF mass analyzers combine the stability of a quadrupole analyzer with the high efficiency, sensitivity, and accuracy of a time-of-flight reflectron mass analyzer and are typically coupled to electrospray ionization sources. The quadrupole can act as a simple quadrupole analyzer to scan across a specified *m/z* range. However, it can also be used to selectively isolate a precursor ion and direct that ion into the collision cell. The resultant fragment ions are then analyzed by the TOF reflectron mass analyzer. Quadrupole-TOF exploits the quadrupole's ability to select a particular ion and the ability of TOF-MS to achieve simultaneous and accurate measurements of ions across the full mass range. Quadrupole-TOF analyzers offer significantly higher sensitivity and accuracy than tandem quadrupole instruments when acquiring full-fragment mass spectra.[59]

FTMS offers high resolution, the ability to perform multiple collision experiments (MSn), and high-accuracy fragment masses (often at the part-per-million level). It is now becoming more common to couple ultrahigh resolution (>105) FTMS to a wide variety of ionization sources, including MALDI, ESI,

APCI, and EI. Quadrupole-FTMS and quadrupole linear ion-trap FTMS mass analyzers that have recently been introduced are typically coupled to electrospray ionization sources. The quad-FTMS combines the stability of a quadrupole analyzer with the high accuracy of an FTMS. A specified *m/z* range can be scanned by using the quadrupole, which can also be used to selectively isolate a precursor ion. This ion can be directed into the collision cell or the FTMS, and the resultant precursor and fragment ions can then be analyzed by the FTMS.[59]

2. Data Mining and Biomarker Selection

The generation of complex data sets in metabolite studies has led to the development of new algorithms[75,76] as well as the incorporation of multivariate statistical-analysis techniques for pattern recognition in biomarker discovery.[5,77–79] This is known as data mining, and can facilitate the discovery of hidden structure in such data. Multivariate techniques can be divided into two classes, known as supervised and unsupervised learning. Supervised learning is when a classification system is given some input data together with the answers, known as the "training set", which can be used to build a model and estimate necessary parameters. Examples of this include artificial neural networks (ANN), self-organizing maps (SOM), and linear discriminant analysis (LDA). Conversely, in unsupervised learning, the algorithm is not given a training set, and so input data is classified in an "unsupervised" manner. The most common unsupervised methods include cluster analysis and PCA.

ANNs are software systems inspired by biological models of the brain. They are capable of learning patterns and relationships from input data and are used to make good pattern-recognition engines and robust classifiers, with an ability to generalize; this enables them to deal with previously unseen and even imprecise input data. ANNs are being used effectively for problems including building nonlinear classification and regression models. Currently, ANNs that can predict patient responses to drugs are being developed; this would enable ideal dosing regimes to be established.[80]

SOMs are a special type of neural net that can be used for visualization, analysis, modeling, clustering, and prediction of high-dimensional data. They can be used to identify nonlinear interactions between parameters and have the ability to learn from new data. This technique has been used to predict the individual outcome of patients receiving 2.5 mg letrozole as second-line endocrine treatment for metastatic breast cancer.[81]

LDA is the most widely used classification technique in several life-science areas, used for classifying samples of unknown classes, such as microarray data.[82,83] However, when analyzing mass-spectrometric data, for which dimensionalities are likely to be larger than training set sizes, the number of features (variables or attributes) must be reduced prior to LDA, for example by using PCA.

Cluster analysis organizes information about variables into a data set, forming relatively homogeneous groups or "clusters" for which the degree of association is strong between data in the same cluster and weak between data in different clusters. Cluster analysis can reveal associations and structure in data that were not previously evident. A combination of proteomics and cluster analysis has been used to accurately classify breast tumor tissues as normal, benign, and cancerous tissues by using just the protein-expression profiles.[84]

PCA simplifies multivariate data by replacing a group of variables with a single new variable.[85] Each new variable is called a principal component (PC), and is a linear combination of the original variables.[86] PCA has been widely used in the reduction of data dimensionality, the investigation of clustering tendency, the detection of outliers, and the visualization of data structure.[87,88] Current uses of PCA include the clustering of gene expression data[89] and the study of proteomic changes associated with neuroblastoma.[90]

Several mass spectrometer manufacturers have their own specific metabolite data-analysis software for use with their instruments, and some software companies have also realized the potential of marketing their own software. However, most companies and research groups involved in metabolite research supplement available data mining techniques with in-house software to enable compound identification and quantification.

Databases

At present, the data collection and analysis steps are by far the most straightforward in metabolite research. A major challenge lies in characterizing interesting metabolites. Although there are well-annotated gene and protein databases that can be searched easily by genomics and proteomics researchers, at present, no such comprehensive tools exist for metabolite researchers.

Current (albeit incomplete) metabolite databases do exist, such as the National Institute of Standards and Technology (NIST) commercial database, which includes mass spectral data for some known metabolites (http://www.nist.gov/srd/nist1.htm). The University of Alberta hosts a mini-library of full mass spectra of newer drugs, metabolites, and some breakdown products, created in May 2001 (http://www.ualberta.ca/~gjones/mslib.htm), which can be downloaded as a zip file. In addition, a new human metabolite database being developed by SRI International called HumanCyc (http://biocyc.org) includes known metabolites and those predicted by algorithms that project metabolic pathways from a genomic sequence, taking into account pathways known to exist in other organisms. A database constructed as part of the Atomic Reconstruction of Metabolism project, compiles metabolite structures together with molecular weight and MS fragmentation data (http://www.metabolome.jp). Another metabolite database being developed is the "tumor metabolome" database, established by Prof. Eigenbrodt and Dr. Mazurek at the Justus Liebig University in Giessen (http://www.metabolic-database.com). Some databases focus purely on electron-impact mass spectrometry data, such as the Wiley Registry of Mass Spectral Data (http://www.wileyregistry.com), the largest commercially available reference library of mass spectra. The GOLM open-access database at the Max Planck Institute of Molecular Plant Physi-

ology (Golm, Germany) also focuses on electron-impact mass spectrometry and is intended as a repository for experiments performed at this institute, as well as for data from collaborators.[91]

To support the management of metabolite data from our own research teams, we have developed METLIN, a web-based data repository (http://metlin.scripps.edu/) to handle and process LC/MS data. METLIN incorporates a diverse collection of spectroscopic and chemical data including LC/MS and high-resolution FTMS spectra, primarily from human biofluids but also some model organisms. The purpose of these data is to aid in metabolite identification through accurate mass measurement and isotopic-pattern evaluation. METLIN also includes an annotated list of known metabolite structural information, both endogenous and drug metabolites, which can be easily cross-correlated with the LC/MS and FTMS data. Further, METLIN provides a number of data-visualization tools including colorized 3D LC/MS plots and histograms. All data are freely available to any interested party. While pharmaceutical companies have developed private internal databases, we believe that this is the first open-access metabolite-profiling data repository.

3. The Crux of Metabolite Profiling: Metabolite Identification

Once potential biomarkers have been selected from metabolite-profiling studies, precise identification is required. This is probably the greatest challenge facing metabolite profiling, given the often limited amount of sample and trace quantities of some metabolites. The use of Q-TOF and FTMS instruments will allow for accurate mass measurements and, subsequently, the identification of biomarkers (Figure 5). In fact, methods for obtaining elemental composition are now typically performed by using high-resolution ESI-FTMS technology for accurate mass determination, together with tandem mass measurements for structural characterization by using CID. Newer hybrid instrument designs are preferable to coupling FTMS/MS to separation techniques such as LC, since MS/MS experiments can be performed outside the magnet.[92] This allows for faster experiments, as high resolution in FTMS is dependent on the presence of high vacuum and by performing MS/MS experiments outside the cell, the ICR cell itself can be maintained at ultra-high vacuum.

Since limited quantities of material are typically available, nanoESI-FTMS will further facilitate these accurate mass measurements (<0.5 ppm) for elemental composition determinations, a critical aspect of metabolite characterization.

However, despite the important information gleaned from mass spectrometry, the lack of comprehensive mass-spectral libraries often limits the identification of molecules from these data alone. Ultimately, the combination of many technologies will be required to identify unknown metabolites in biofluids. Other technologies that can be utilized in metabolite identification are high-sensitivity capillary NMR experiments, which provide metabolite-structure characterization down to the low microgram level.[93,94] In addition, chemical-modification experiments also offer structural information.

A recent example of this approach to metabolite profiling and characterization allowed for the identification of a completely novel set of molecules as taurine-conjugated fatty acids.[11] The initial strategy was to use LC/MS to compare the nervous system metabolomes of wild-type and fatty acid amide hydrolase (FAAH) enzyme-inactivated organisms. However, the relative hydrolytic activity that FAAH exhibited for lipid metabolites in vitro was not predictive of the identity of specific FAAH substrates in vivo. Thus, a three-step method was used to identify these metabolites, 1) ultra high accuracy FTMS mass measurements, 2) high-accuracy tandem mass analysis with a Q-TOF, and 3) chemical synthesis of potential candidates based on the results of (1) and (2). An example of the data is shown in Figure 6. Accurate mass measurements of the $m/z = 446$ metabolite by ESI-FTMS provided an exact mass of 446.3310, which corresponded to a molecular formula of $C_{24}H_{48}NO_4S$. The theoretical and experimental isotope patterns for $C_{24}H_{48}NO_4S$ overlaid well, including the splitting pattern of the $[M+2]$ isotope caused by differences in the mass between two ^{13}C isotopes and one ^{34}S isotope. Similar spectra were obtained for the m/z 472 and 474 metabolites, providing exact masses of $C_{26}H_{50}NO_4S$ (472.3466) and $C_{26}H_{52}NO_4S$ (474.3623), respectively. MS/MS analyses of the natural and synthetic metabolites using a Micromass Q-TOF instrument led to their structural assignments as *N*-acyl taurines.

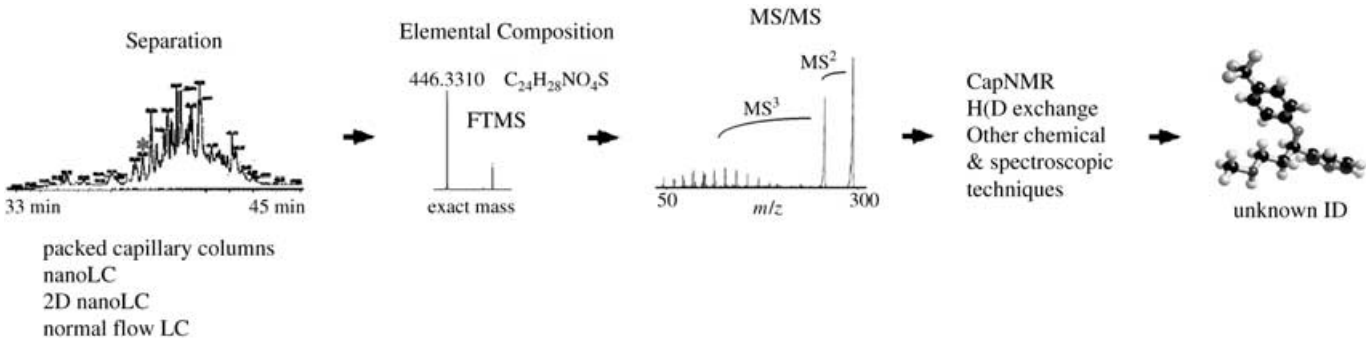

Figure 5. Diagrammatic representation of an approach to profiling and characterizing small-molecule biomarkers from biofluids. Once one or more biomarkers of unknown identity have been selected, characterization can be facilitated through accurate mass measurements and tandem mass-spectral data. Other spectroscopic techniques, such as NMR are very useful for the complete identification of unknown compounds.

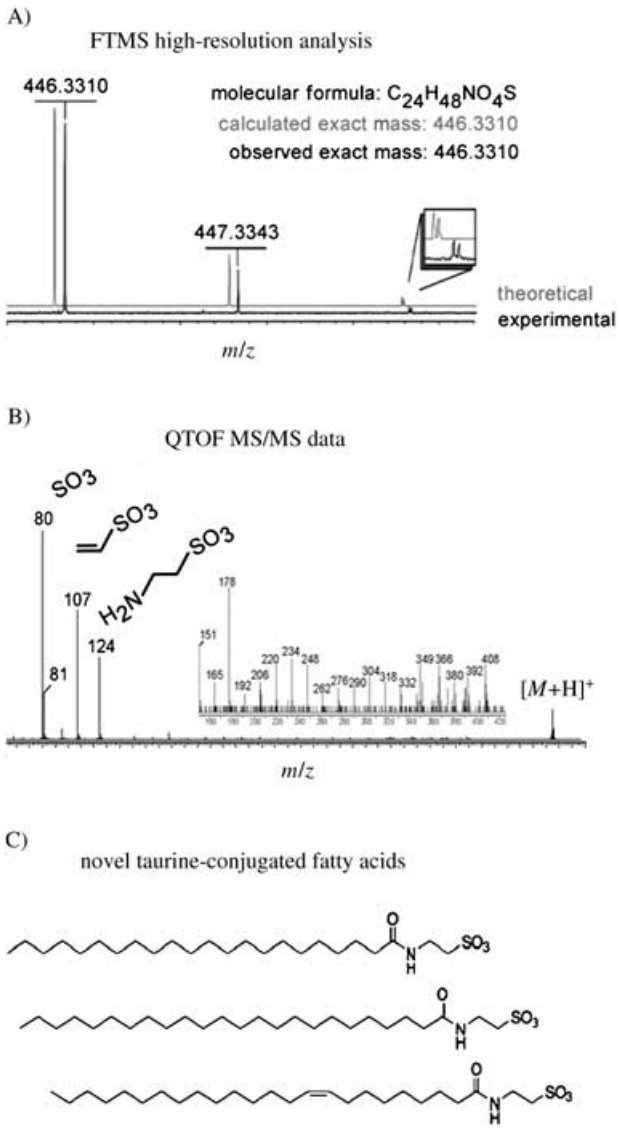

Figure 6. Chemical characterization of an unknown class of brain metabolites as *N*-acyl taurines. A) Analysis of the $m/z = 446$ metabolite by ESI-FTMS (Bruker APEX III 7T instrument) gave an exact mass of 446.3310, which corresponds to $C_{24}H_{48}NO_4S$. The calculated isotope pattern for $C_{24}H_{48}NO_4S$ (gray) overlaid well with the experimental spectrum, including the splitting pattern of the $[M+2]$ isotope (inset). Similar spectra were obtained for the $m/z =$ 472 and 474 metabolites to give $C_{26}H_{50}NO_4S$ (472.3466) and $C_{26}H_{52}NO_4S$ (474.3623), respectively. B) MS/MS data of an *N*-acyl taurine obtained on a Micromass Q-TOF instrument. Prominent fragments corresponding to taurine (124), vinylsulfonic acid (107), and sulfur trioxide (80) are highlighted, as well as a pattern of progressive loss of 14 mass units from $m/z = 150$–430; this is indicative of a fatty acyl chain (inset). C) Structures of the three main *N*-acyl taurines identified.

4. Applications of Metabolite Profiling Research

There is a wide range of applications for metabolite research in areas such as drug design and discovery, health and disease, nutrition, plant biotechnology, and microbiology.[95] Metabolite profiling can be used to determine drug efficacy and/or toxicity, to diagnose or predict disease states, or to classify patient groups based on their specific metabolism. By studying alterations in metabolites brought on by disease or drug intervention, their relationships to changes in gene expression, protein expression, and enzyme activity can be explored.[96]

Metabolite profiling is already promising to bring a unique and complementary perspective to biomedical and therapeutic research, and is thus evolving into an important part of the drug-discovery and -development process. The measurement of alterations in metabolite concentrations can help to determine the range of biochemical effects induced by a therapeutic agent. The effect of a potential drug on an animal can be studied by measuring the changes in endogenous metabolite levels over a time course.[97] Multivariate statistical-analysis techniques can be used to highlight differences between dosed and control samples, an approach that is now being investigated by large pharmaceutical companies to screen potential drug compounds for toxicity and lead-compound selection, as well as investigating in vivo efficacy in animals. In this way, metabolite profiling might dramatically reduce the costs of drug development, by eliminating the progression of compounds that would eventually fail due to toxicity. In the development phase, metabolite profiling could also aid in the discovery of new preclinical and clinical safety biomarkers. Specific metabolites are already measured in biofluids in order to identify drugs with liver or kidney toxicity—it is known that succinate, glycine, and dimethylamine in the blood indicate kidney damage.[25] Further, metabolite-profiling techniques might be able to highlight the responses of different groups of patients to a given drug. These types of investigations have already been carried out in diabetic mice.[98]

Even when there is no therapeutic intervention, examining sequentially collected plasma or urine samples over a particular time course would enable endogenous metabolite-profile changes to be examined.[99, 100] Furthermore, the timing of the appearance of small-molecule markers in particular biofluids could be studied. For example, the combination of methods such as NMR and pattern recognition (NMR-PR) has been used to investigate the hormonal cycle of rats,[101] highlighting the utility of such techniques in the investigation of physiological rhythms and variation.

Metabolite profiling can be used to investigate the effect of nutrition on metabolite concentrations in the body and to correlate these metabolite concentrations with disease. A widely used single-metabolite marker is the level of serum cholesterol as an indicator for the increased risk of heart disease.[102] Newer research has aimed to investigate the entire plasma small-molecule profile, which reflects the biochemical status of an individual at a particular point in time.[103] However, links need to be made between metabolism in different areas of the body, such as tissue and plasma, in order to fully understand this complex area. Studies showing that selected serum metabolites can distinguish between groups of rats fed diets containing very different calorie levels[104, 105] have implications for understanding the effect of nutrition on serum metabolite profiles.

As the value and throughput of plant biotechnology is increasing, so is the use of mass spectrometry in metabolite profiling, as well as bioinformatics to interpret the vast amount of data that is generated. It is estimated that there could be up to 200 000 different metabolites in the plant kingdom,[106] and the number of these compounds that performs precise biological functions is unknown. Furthermore, the convoluted nature of plant metabolism complicates interpretation of metabolite-profiling data. However, recent advances in laboratory technologies have allowed the quantitation of over 1000 small molecules in a single leaf extract, as well as more than 500 compounds from potato tubers.[53,107] The combination of mass spectrometry and sophisticated data-analysis techniques has increased the already numerous applications of metabolite profiling in the area of plant biotechnology, which includes the production of pharmaceuticals in plants,[108] thus decreasing the need for pesticide or fertilizer application and enhancing the nutritional value of food.

Another application of metabolite profiling is the typing and characterization of microorganisms.[109,110] Cellular proteins or metabolites specific to an individual microorganism are determined by mass-spectral analysis of the intact cellular suspensions, thus enabling organisms present in unknown samples to be identified. This approach has applications in the biological, medical, and pharmaceutical arenas, where samples from health institutions, industrial sites, agricultural products, and food processing can be monitored for suspected bacterial contamination rapidly and cost-effectively.

Summary and Outlook

As with PKU, many metabolic disorders can now be easily monitored, and many more are likely to be discovered either related to disease or as a function of an administered drug. The area of metabolite profiling is expanding rapidly, and the applications for this science range from basic biochemistry to clinical-biomarker discovery. The primary challenge in metabolite profiling is in the generation of a comprehensive, quantitative profile of the thousands of components present in biofluids, an issue that is largely being addressed with new LC/MS technology. A second challenge is data analysis, here the development of sophisticated bioinformatics software packages will ultimately drive the discovery process. A third challenge is in structurally characterizing physiologically important molecules observed in the LC/MS experiments. Given these challenges, it is encouraging that new biomarkers for major diseases such as atherosclerosis, muscular dystrophy, and various cancers have already been identified. Ultimately it is the discovery of novel metabolites[10,11] and correlating the changes of multiple metabolites with physiological events that make this area alluring and challenging.

Acknowledgements

The authors would like to thank Colin Smith of the Scripps Research Institute for his implementation of and involvement with the METLIN Database and subsequent data analysis tools. We would also like to acknowledge NIH grants GM055775 and MH62261 for supporting E.W. and G.S.

Keywords: biomarker characterization · liquid chromatography · mass spectrometry · metabolite database · metabolite profiling

[1] G. J. Dear, J. Ayrton, R. Plumb, I. J. Fraser, *Rapid Commun. Mass Spectrom.* **1999**, *13*, 456–463.
[2] N. Zhang, S. T. Fountain, H. Bi, D. T. Rossi, *Anal. Chem.* **2000**, *72*, 800–806.
[3] J. P. Shockcor, E. Holmes, *Curr. Top. Med. Chem.* **2002**, *2*, 35–51.
[4] P. R. Tiller, L. A Romanyshyn, *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1225–1231.
[5] R. S. Plumb, C. L. Stumpf, J. H. Granger, J. Castro-Perez, J. N. Haselden, G. J. Dear, *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2632–2638.
[6] Y. Deng, J. T. Wu, H. Zhang, T. V. Olah, *Rapid Commun. Mass Spectrom.* **2004**, *18*, 1681–1685.
[7] H. H. Maurer, *Anal. Bioanal. Chem.* **2005**, *381*, 110–118.
[8] R. F. Staack, E. Varesio, G. Hopfgartner, *Rapid Commun. Mass Spectrom.* **2005**, *19*, 618–626.
[9] D. Q. Liu, C. E. Hop, *J. Pharm. Biomed. Anal.* **2005**, *37*, 1–18.
[10] B. F. Cravatt, O. Prospero-Garcia, G. Siuzdak, N. B. Gilula, S. J. Henriksen, D. L. Boger, R. A. Lerner, *Science* **1995**, *268*, 1506–1509.
[11] A. Saghatelian, S. A. Trauger, E. J. Want, E. G. Hawkins, G. Siuzdak, B. F. Cravatt, *Biochemistry* **2004**, *43*, 14332–14339.
[12] D. H. Chace, D. S. Millington, N. Terada, S. G. Kahler, C. R. Roe, L. F. Hofman, *Clin. Chem.* **1993**, *39*, 66–71.
[13] D. H. Chace, S. L. Hillman, D. S. Millington, S. G. Kahler, C. R. Roe, E. W. Naylor, *Clin. Chem.* **1995**, *41*, 62–68.
[14] D. H. Chace, S. L. Hillman, D. S. Millington, S. G. Kahler, B. W. Adam, H. L. Levy, *Clin. Chem.* **1996**, *42*, 349–355.
[15] D. H. Chace, B. W. Adam, S. J. Smith, J. R. Alexander, S. L. Hillman, W. H. Hannon, *Clin. Chem.* **1999**, *45*, 1269–1277.
[16] D. H. Chace, *Chem. Rev.* **2001**, *101*, 445–477.
[17] D. H. Chace, J. C. DiPerna, B. L. Mitchell, B. Sgroi, L. F. Hofman, E. W. Naylor, *Clin. Chem.* **2001**, *47*, 1166–1182.
[18] D. H. Chace, T. A. Kalas, E. W. Naylor, *Annu. Rev. Genomics Hum. Genet.* **2002**, *3*, 17–45.
[19] D. H. Chace, R. Pons, C. A. Chiriboga, D. J. McMahon, I. Tein, E. W. Naylor, D. C. De Vivo, *Pediatr. Res.* **2003**, *53*, 823–829.
[20] D. H. Chace, *Clin. Chem.* **2003**, *49*, 1227–1229.
[21] D. H. Chace, T. A. Kalas, E. W. Naylor, *Clin. Chem.* **2003**, *49*, 1797–1817.
[22] C. R. Scriver in *The Metabolic and Molecular Bases of Inherited Disease*, 7th ed., McGraw–Hill, New York, **1995**.
[23] T. Niwa, *Clin. Chim. Acta* **1995**, *241–242*, 75–152.
[24] J. C. Lindon, J. K. Nicholson, E. Holmes, H. Antti, M. E. Bollard, H. Keun, O. Beckonert, T. M. Ebbels, M. D. Reily, D. Robertson, G. J. Stevens, P. Luke, A. P. Breau, G. H. Cantor, R. H. Bible, U. Niederhauser, H. Senn, G. Schlotterbeck, U. G. Sidelmann, S. M. Laursen, A. Tymiak, B. D. Car, L. Lehman-McKeeman, J. M. Colet, A. Loukaci, C. Thomas. *Toxicol. Appl. Pharmacol.* **2003**, *187*, 137–146.
[25] J. K. Nicholson, J. Connelly, J. C. Lindon, E. Holmes, *Nat. Rev. Drug Discovery* **2002**, *1*, 153–161.
[26] N. V. Reo, *Drug Chem. Toxicol.* **2002**, *25*, 375–382.
[27] J. L. Griffin, *Curr. Opin. Chem. Biol.* **2003**, *7*, 648–654.
[28] R. Bligny, R. Douce, *Curr. Opin. Plant Biol.* **2001**, *4*, 191–196.
[29] R. G. Ratcliffe, Y. Shachar-Hill, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2001**, *52*, 499–526.
[30] J. L. Griffin, H. J. Williams, E. Sang, K. Clarke, C. Rae, J. K. Nicholson, *Anal. Biochem.* **2001**, *293*, 16–21.
[31] K. S. Solanky, N. J. Bailey, E. Holmes, J. C. Lindon, A. L. Davis, T. P. Mulder, J. P. Van Duynhoven, J. K. Nicholson, *J. Agric. Food Chem.* **2003**, *51*, 4139–4145.
[32] J. T. Brindle, J. K. Nicholson, P. M. Schofield, D. J. Grainger, E. Holmes, *Analyst* **2003**, *128*, 32–36.

[33] M. Coen, E. M. Lenz, J. K. Nicholson, I. D. Wilson, F. Pognan, J. C. Lindon, *Chem. Res. Toxicol.* **2003**, *16*, 295 – 303.
[34] T. Niwa, *J. Chromatogr.* **1986**, *379*, 313 – 345.
[35] E. Gelpi, *J. Chromatogr. A* **1995**, *703*, 59 – 80.
[36] C. Deng, Y. Deng, B. Wang, X. Yang, *J. Chromatogr. B* **2002**, *780*, 407 – 413.
[37] C. Deng, C. Shang, Y. Hu, X. Zhang, *J. Chromatogr. B* **2002**, *775*, 115 – 120.
[38] C. Deng, Y. Deng, *J. Chromatogr. B* **2003**, *792*, 261 – 268.
[39] C. Deng, N. Li, X. Zhang, *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2558 – 2564.
[40] D. B. Ramsden, M. Farmer, M. N. Mohammed, P. Willetts, *Biomed. Mass Spectrom.* **1984**, *11*, 193 – 198.
[41] L. M. Thienpont, V. I. De Brabandere, D. Stockl, A. P. De Leenheer, *Biol. Mass Spectrom.* **1994**, *23*, 475 – 482.
[42] V. I. De Brabandere, P. Hou, D. Stockl, L. M. Thienpont, A. P. De Leenheer, *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1099 – 1103.
[43] J. Sjövall, A. M. Lawson, K. D. Setchell, *Methods Enzymol.* **1985**, *111*, 63 – 113.
[44] S. Perwaiz, D. Mignault, B. Tuchweber, I. M. Yousef, *Lipids* **2002**, *37*, 1093 – 1100.
[45] C. H. Shackleton, *J. Chromatogr.* **1986**, *379*, 91 – 156.
[46] M. P. Caulfield, T. Lynn, M. E. Gottschalk, K. L. Jones, N. F. Taylor, E. M. Malunowicz, C. H. Shackleton, R. E. Reitz, D. A. Fisher, *J. Clin. Endocrinol. Metab.* **2002**, *87*, 3682 – 3690.
[47] S. A. Wudy, M. F. Hartmann, *Horm. Metab. Res.* **2004**, *36*, 415 – 422.
[48] D. T. Forman, *Ann. Clin. Lab. Sci.* **1991**, *21*, 85 – 93.
[49] P. M. Jones, A. B. Burlina, M. J. Bennett, *J. Inherited Metab. Dis.* **2000**, *23*, 745 – 750.
[50] P. M. Jones, R. Quinn, P. V. Fennessey, S. Tjoa, S. I. Goodman, S. Fiore, A. B. Burlina, P. Rinaldo, R. L. Boriack, M. J. Bennett, *Clin. Chem.* **2000**, *46*, 149 – 155.
[51] C. H. Shackleton, *Endocr. Rev.* **1985**, *6*, 441 – 486.
[52] C. H. Shackleton, E. Roitman, L. Kratz, R. Kelley, *Prenatal Diagn.* **2001**, *21*, 207 – 212.
[53] O. Fiehn, J. Kopka, P. Dormann, T. Altmann, R. N. Trethewey, L. Willmitzer, *Nat. Biotechnol.* **2000**, *18*, 1157 – 1161.
[54] A. Goossens, S. T. Hakkinen, I. Laakso, T. Seppanen-Laakso, S. Biondi, V. De Sutter, F. Lammertyn, A. M. Nuutila, H. Soderlund, M. Zabeau, D. Inze, K. M. Oksman-Caldentey, *Proc. Natl. Acad. Sci.* **2003**, *100*, 8595 – 8600.
[55] L. W. Sumner, P. Mendes, R. A. Dixon, *Phytochemistry* **2003**, *62*, 817 – 836.
[56] A. P. Dias, J. Brown, P. Bonello, E. Grotewold, *Methods Mol. Biol.* **2003**, *236*, 415 – 426.
[57] J. Kopka, A. Fernie, W. Weckwerth, Y. Gibon, M. Stitt, *Adv. Genome Biol.* **2004**, *5*, 109.
[58] M. M. Kushnir, G. Komaromy-Hiller, B. Shushan, F. M. Urry, W. L. Roberts, *Clin. Chem.* **2001**, *47*, 1993 – 2002.
[59] G. Siuzdak, *The Expanding Role of Mass Spectrometry in Biotechnology*, MCC Press, San Diego, CA, **2003**.
[60] H. L. Levy, *Clin. Chem.* **1998**, *44*, 2401 – 2402.
[61] B. K. Matuszewski, M. L. Constanzer, C. M. Chavez-Eng, *Anal. Chem.* **1998**, *70*, 882 – 889.
[62] E. T. Gangl, M. M. Annan, N. Spooner, P. Vouros, *Anal. Chem.* **2001**, *73*, 5635 – 5644.
[63] S. A. Gustavsson, J. Samskog, K. E Markides, B. Långström, *J. Chromatogr. A* **2001**, *37*, 41 – 47.
[64] R. S. Plumb, H. Warwick, D. Higton, G. J. Dear, D. N. Mallett, *Rapid Commun. Mass Spectrom.* **2001**, *15*, 297 – 303.
[65] H. J. Bremer in *Disturbances of Amino Acid Metabolism: Clinical Chemistry and Diagnosis*, Urban and Schwarzenberg, Baltimore, **1981**.
[66] F. A. Hommes in *Techniques in Diagnostic Human Biochemical Genetics: A Laboratory Manual*, Wiley-Liss, New York, **1991**.
[67] D. Chelius, T. Zhang, G. Wang, R. F. Shen, *Anal. Chem.* **2003**, *75*(23), 6658 – 6665.
[68] E. Nagele, M. Vollmer, P. Horth, *J. Chromatogr. A* **2003**, *1009*, 197 – 205.
[69] K. Chatman, T. Hollenbeck, L. Hagey, M. Vallee, R. Purdy, F. Weiss, G. Siuzdak, *Anal. Chem.* **1999**, *71*, 2358 – 2363.
[70] W. J. Griffiths, S. Liu, Y. Yang, R. H. Purdy, J. Sjövall, *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1591 – 1610.
[71] J. Abian, A. J. Oosterkamp, E. Gelpí, *J. Mass Spectrom.* **1999**, *34*, 244 – 254.
[72] J. I. Castrillo, A. Hayes, S. Mohammed, S. J. Gaskell, S. G. Oliver, *Phytochemistry* **2003**, *62*, 929 – 937.
[73] W. C. Byrdwell, *Lipids* **2001**, *36*, 327 – 346.
[74] N. De Marchi, L. De Petrocellis, P. Orlando, F. Daniele, F. Fezza, V. Di Marzo, *Lipids Health Dis.* **2003**, *2*, 5.
[75] C. A. Hastings, S. M. Norto, S. Roy, *Rapid Commun. Mass Spectrom.* **2002**, *16*, 462 – 467.
[76] A. Floter, J. Nicolas, T. Schaub, J. Selbig, *Bioinformatics* **2004**, *20*, 1491 – 1494.
[77] S. M. Norton, P. Huyn, C. A. Hastings, J. C. Heller, *Curr. Opin. Drug Discov. Devel.* **2001**, *4*, 325 – 331.
[78] H. Idborg, P. O. Edlund, S. P. Jacobsson, *Rapid Commun. Mass Spectrom.* **2004**, *18*, 944 – 954.
[79] P. Jonsson, J. Gullberg, A. Nordström, M. Kusano, M. Kowalczyk, M. Sjöström, T. Moritz, *Anal. Chem.* **2004**, *76*, 1738 – 1745.
[80] A. E. Gaweda, A. A. Jacobs, M. E. Brier, J. M. Zurada, *Neural. Networks* **2003**, *16*, 841 – 845.
[81] M. Wischnewsky, P. Schmid, F. Hegemann, K. Possinger, *Ann. Oncol.* **2000**, *11 (Suppl.4)*, 25.
[82] Y. Hakak, J. R. Walker, C. Li, W. H. Wong, K. L. Davis, J. D. Buxbaum, V. Haroutunian, A. A. Fienberg, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4746 – 4751.
[83] S. Dudoit, J. Fridlyand, T. P. Speed, *J. Am. Stat. Assoc.* **2002**, *97*, 77 – 87.
[84] M. V. Dwek, A. A. Alaiya, *Br. J. Cancer* **2003**, *89*, 305 – 307.
[85] I. T. Jolliffe in *Principal Component Analysis*, Springer, New York, **1986**.
[86] R. A. Calvo, M. Partridge, M. Jabri in *Proceedings of the Ninth Australian Conference on Neural Networks*, Brisbane, **1998**.
[87] H. Martens, T. Naes in *Multivariate Calibration*, Wiley, Chicester, **1989**.
[88] E. R. Malinowski in *Factor Analysis in Chemistry*, 2nd ed., Wiley, New York, **1991**.
[89] K. Y. Yeung, W. L. Russo, *Bioinformatics* **2001**, *17*, 763 – 774.
[90] E. Marengo, E. Robotti, P. G. Righetti, N. Campostrini, J. Pascali, M. Ponzoni, M. Hamdan, H. Astner, *Clin. Chim. Acta* **2004**, *345*, 55 – 67.
[91] J. Kopka, N. Schauer, S. Krueger, C. Birkemeyer, B. Usadel, E. Bergmuller, P. Dormann, Y. Gibon, M. Stitt, L. Willmitzer, A. R. Fernie, D. Steinhauser, *Bioinformatics* **2005**, *21*, 1635 – 1638.
[92] S. M. Patrie, J. P. Charlebois, D. Whipple, N. L. Kelleher, C. L. Hendrickson, J. P. Quinn, A. G. Marshall, B. Mukhopadhyay, *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1099 – 1108.
[93] G. Schlotterbeck, A. Ross, R. Hochstrasser, H. Senn, T. Kühn, D. Marek, O. Schett, *Anal. Chem.* **2002**, *74*, 4464 – 4471.
[94] D. L. Olson, J. A. Norcross, M. O'Neil-Johnson, P. F. Molitor, D. J. Detlefsen, A. G. Wilson, T. L. Peck, *Anal. Chem.* **2004**, *76*, 2966 – 2974.
[95] A. Buchholz, J. Hurlebaus, C. Wandrey, R. Takors, *Biomol. Eng.* **2002**, *19*, 5 – 15.
[96] E. Holmes, A. W. Nicholls, J. C. Lindon, S. C. Connor, J. C. Connelly, J. N. Haselden, S. J. Damment, M. Spraul, P. Neidig, J. K. Nicholson, *Chem. Res. Toxicol.* **2000**, *13*, 471 – 478.
[97] 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.* **2002**, *16*, 1991 – 1996.
[98] S. M. Watkins, J. B. German, *Curr. Opin. Mol. Ther.* **2002**, *4*, 224 – 228.
[99] E. Holmes, A. W. Nicholls, J. C. Lindon, S. Ramos, M. Spraul, P. Neidig, S. C. Connor, J. Connelly, S. J. Damment, J. Haselden, J. K. Nicholson, *NMR Biomed.* **1998**, *11*, 235 – 244.
[100] J. K. Nicholson, J. C. Lindon, E, Holmes, *Xenobiotica* **1999**, *29*, 1181 – 1189.
[101] M. E. Bollard, E. Holmes, J. C. Lindon, S. C. Mitchell, D. Branstetter, W. Zhang, J. K. Nicholson, *Anal. Biochem.* **2001**, *295*, 194 – 202.
[102] Z. Chen, R. Peto, R. Collins, S. MacMahon, J. Lu, W. Li, *Br. Med. J.* **1991**, *303*, 276 – 282.
[103] S. M. Watkins, J. B. German, *Curr. Opin. Biotechnol.* **2002**, *13*, 512 – 516.
[104] K. E. Vigneau-Callahan, A. I. Shestopalov, P. E. Milbury, W. R. Matson, B. S. Kristal, *J. Nutr.* **2001**, *131*, 924S – 932S.
[105] H. Shi, K. E. Vigneau-Callahan, A. I. Shestopalov, P. E. Milbury, W. R. Matson, B. S. Krista, *J. Nutr.* **2002**, *132*, 1039 – 1046.
[106] O. Fiehn, *Plant Mol. Biol.* **2002**, *48*, 155 – 171.
[107] U. Roessner, C. Wagner, J. Kopka, R. N. Trethewey, L. Willmitzer, *Plant J.* **2000**, *23*, 131 – 142.

[108] G. Giddings, G. Allison, D. Brooks, A. Carter, *Nat. Biotechnol.* **2000**, *18*, 1151 – 1155.
[109] T. Krishnamurthy, M. T. Davis, D. C. Stahl, T. D. Lee, *Rapid Commun. Mass Spectrom.* **1999**, *13*, 39 – 49.
[110] F. Xiang, G. A. Anderson, T. D. Veenstra, M. S. Lipton, R. D. Smith, *Anal. Chem.* **2000**, *72*, 2475 – 2481.

Received: April 8, 2005
Published online on October 4, 2005

