Published in final edited form as:

Trends Analyt Chem. 2007 June 1; 26(6): 625–636. doi:10.1016/j.trac.2007.02.009.

Highly-parallel metabolomics approaches using LC-MS² for pharmaceutical and environmental analysis

Sunil Bajad and Vladimir Shulaev*

Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

Abstract

The ‘omics’ approaches – genomics, proteomics and metabolomics – are based on high-throughput, high-information-content analysis. Using these approaches, as opposed to targeting one or a few analytes, a holistic understanding of the composition of a sample can be obtained. These approaches have revolutionized sample-analysis and data-processing protocols. In metabolomic studies, hundreds of small molecules are simultaneously analyzed using analytical platforms (e.g., gas chromatography-mass spectrometry (GC-MS) or liquid chromatography coupled to tandem mass spectrometry (LC-MS²)). This philosophy of holistic analysis and the application of high-throughput, high-information-content analysis offer several advantages. In this article, we compare the conventional analytical approach of one or a few analytes per sample to the LC-MS²-based metabolomics-type approach in the context of pharmaceutical and environmental analysis.

Keywords

Chromatography; Environmental; LC-MS²; Metabolomics; Pharmaceutical

1. Introduction

Genomics and proteomics studies have become a common tool in many research areas. In recent years, a new ‘omics’ discipline – metabolomics – has been gaining popularity due to its applications in diverse fields (e.g., functional genomics, drug discovery, toxicology, environmental science, nutritional science and disease diagnosis) [1–4]. In addition to their significance in understanding biology, the advances in ‘omics’ technologies have changed the way we now obtain, analyze and present data from biological samples. These approaches have significantly increased the rate of sample screening as well as the information content in the data obtained. These holistic approaches, applied together or individually, can provide novel biological insights that can often be missed with conventional highly-targeted sample-analysis approaches.

Currently, holistic analysis is mainly restricted to biochemical research. However, no matter the discipline, the advantages of screening a sample for hundreds of molecules simultaneously are enormous. Hence, the concept needs to be tested and applied to other fields, such as pharmaceutical and environmental analysis.

*Corresponding author. Tel.: +1 (540) 231-3489, Fax: +1 (540) 231-2606, E-mail: vshulaev@vbi.vt.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Though the analytical platforms used for genomics and proteomics are relatively specific, metabolomics uses very common analytical platforms such as gas chromatography with mass spectrometry (GC-MS), liquid chromatography with MS (LC-MS) and nuclear magnetic resonance (NMR) spectroscopy. Hence, the application of metabolomics to pharmaceutical and environmental analyses, which utilize the same analytical platform, should be relatively easy and would provide several advantages (Table 1). An excellent example of the application of metabolomics-type study to pharmaceutical analysis is the LC with tandem MS (LC-MS²) method developed by Kirchherr et al. [5] for simultaneous analysis of 48 drugs in human serum.

We provide a critical overview of the metabolomics approach and its potential for pharmaceutical and environmental analysis. The metabolomics approach discussed here is related to the determination of the pharmaceuticals and environmental chemicals in a sample and should not be confused with, for example, environmental metabolomics, which deals with the effect of environment on the biochemical processes in organisms. Thus, the metabolomics applications to analyze the pharmaceutical composition of a sample and environmental metabolomics differ but can be complementary.

2. A highly-parallel analytical approach

Metabolomics and its applications in various research areas have been extensively reviewed [6,7]. Here, we provide just a brief overview in order to lay down a foundation for our discussion of applications of metabolomics approaches to pharmaceutical and environmental analyses. Metabolomics involves analysis of all or a large number of cellular metabolites. It is usually used to compare the metabolite levels in organisms under a given condition (e.g., comparison of mutants to wild-type strains). A typical metabolomics study involves

1. sample processing for unbiased metabolite extraction;
2. simultaneous analysis of a sample for a large number of small molecules using NMR or MS (with or without prior chromatographic separation);
3. data pre-processing for background subtraction, alignment and normalization; and,
4. multivariate statistical analysis to compare samples with regard to their metabolite content and levels [8].

Sample preparation for metabolomics analysis requires unbiased extraction of metabolites so that those with a wide range of physicochemical properties are detected and quantitated. For determination of metabolite levels two major approaches are commonly used:

1. targeted analysis; and,
2. non-targeted analysis.

As the name indicates, the targeted analysis involves quantitation of a known set of metabolites whose chemical identities are known. A well-defined method is required for performing targeted analysis. This approach is based on a specific selected reaction monitoring (SRM) developed for each targeted metabolite. Hundreds of these SRMs are then monitored by MS in a single separation.

On the other hand, non-targeted analysis involves analysis of metabolites without prior knowledge of their identities.

Both approaches have advantages and disadvantages which will be discussed in the following sections. For both approaches, several analytical platforms such as NMR, GC-MS, LC-MS and capillary electrophoresis-MS (CE-MS) are used. However, LC-MS is the most commonly

used, due to its sensitivity and amenability for detecting highly polar molecules (e.g., nucleotides) as well as its ability to provide structural information.

The raw data obtained in metabolomic studies normally require pre-processing before being subjected to statistical analysis. Software tools are used to correct the data for background noise, peak alignment and normalization. The corrected data are then subjected to supervised and non-supervised multivariate statistical analysis (e.g., Analysis of Variance (ANOVA), Principle Component Analysis (PCA), hierarchical clustering (HC), self-organized maps (SOMs), partial least squares (PLS), and Genetic Algorithm-Discriminant Function Analysis (GA-DFA)).

3. Applications to pharmaceutical and environmental analysis

The main objective of metabolomic studies is to identify the differences in metabolite content of biological samples. Typically, the relative quantitation of a large number of metabolites belonging to diverse chemical classes (e.g., amino acids, carbohydrates, carboxylic acids, nucleotides and their precursors and co-factors, such as glutathione) is achieved from a single sample. Similarly, in pharmaceutical and environmental studies, complex matrices containing diverse chemical components (e.g., drugs and their metabolites, toxic chemicals, and pesticides) are analyzed. Hence, performing metabolomics-type analysis of these complex matrices would present significant challenges for sample preparation, analysis, data processing and interpretation. We will discuss each of the sample-processing steps in the context of their applications to pharmaceutical and environmental analysis.

3.1. Sample preparation

There are several differences in sample collection and processing that have to be considered for conventional versus metabolomics-type analysis (Table 2).

A metabolomics study requires rapid quenching of a sample to stop the biochemical processes, in order to obtain a snapshot of the biochemical composition at the time of sampling. For this purpose, usually cold (-20°C or less) organic solvent (e.g., buffered/unbuffered methanol or acetonitrile) is added to the sample [9]. Next, the stability of a sample obtained is of the utmost importance, as several molecules (e.g., nucleotides and thiols) are very unstable. A metabolomics-type approach for pharmaceutical and environmental analysis would require careful consideration of the nature of matrix and the analytes of interest. Microbes are commonly present in the environmental samples and are capable of degrading or metabolizing organic chemicals including proteins [10]. Hence, filtration followed by storage at -20°C to -80°C or immediate analysis might be required. Quenching the sample with cold (-20°C or lower) organic solvent after the filtration should also be considered, as enzymes from dead microorganisms might be present in the sample. The presence of reactive components (e.g., oxidizing agents, reducing agents or metals) may change the composition of the sample after collection. Proper handling of the sample is therefore necessary to avoid sample-to-sample variability and the loss of unstable and volatile components. Proper sample collection and handling is critical in a clinical environment, given that blood and urine samples contain numerous reactive components, including drug-metabolizing enzymes. When analyzing several analytes together, the stability of each individual analyte and its reactivity in the mixture have to be considered, and steps should be taken to protect it from transformation or degradation. For example, as reported by Kirchherr et al. [5], simultaneous analysis of 48 antidepressant and antipsychotic drugs in human serum required protection of the sample from light during collection, processing and storage, as some of the drugs were light sensitive. A great deal of care was also required during transport of the samples from collection site to the analysis site, as is often necessary for environmental studies.

Direct analysis of a sample without any extraction can be tried in some cases to avoid loss of any analytes during extraction [11]. However, such an approach is not suitable for samples containing proteins, and matrix effects might distort the analytical make-up of the sample.

Sample-processing techniques that can be used for simultaneous analysis of large number of pharmaceuticals and environmental chemicals are summarized in Table 3.

For plasma and serum samples, protein precipitation with organic solvent might work well, as demonstrated for the analysis of a large number of pharmaceuticals in human plasma [5]. Solid-phase extraction (SPE) [12] or liquid-liquid extraction (LLE) [13] might be a good choice if the analytes of interest have similar physicochemical properties. However, a simple one-step extraction may not be possible if targeted compounds involve polar, non-polar and neutral compounds. These cases might require a complex extraction (e.g., sequential SPE) [14].

Considering that there is no universal extraction method, a compromise is needed between the number of molecules to be extracted versus efficiency and reproducibility of extraction. The suitability of the extraction solvent for the MS analysis also has to be taken into account.

3.2. Sample analysis

An analytical platform capable of detecting a large number of molecules with diverse chemical functionalities is required for metabolomics analysis. It must combine high sensitivity with a wide dynamic range. Several analytical techniques routinely used in metabolomics studies can be applied to pharmaceutical and environmental analysis, including: NMR [15]; direct infusion MS (diMS) [16]; CE-MS [17]; GC-MS or LC-MS [18,19]; and, LC with electrochemical detection (LC-ECD) [20]. Each of these techniques can simultaneously detect hundreds of molecules in a sample and has certain advantages and disadvantages [21].

Careful selection of the platform is required, depending upon the nature of the sample and the targeted set of metabolites (e.g., NMR may not be a good choice for low-abundance analytes, while CE or LC combined with high-resolution MS (HRMS) would be the platform of choice if the sample was expected to contain a number of isobaric or isomeric compounds. Over the years, LC-MS has emerged as the popular option for metabolomics due to several advantages (e.g., selectivity, sensitivity, versatility and the ability to identify unknown analytes) [22]. Due to these features, this technique may be the obvious choice for metabolomics-type pharmaceutical and environmental analysis. In the following sections, we will review the LC-MS-based approaches for highly-parallel analysis.

3.2.1. Separation—Though direct infusion of samples into a mass spectrometer without prior separation has been tried for analysis of hundreds of molecules [23], the simultaneous presence of many analytes in the MS source usually causes ion suppression. With this approach, many low-abundance analytes may not be detected due to competition or ion suppression caused by high-abundance molecules. The purpose of the chromatographic separation in metabolomics is to reduce analyte crowding in the MS source and to separate as many analytes as possible, including isomeric or isobaric analytes, using a single analytical method.

So far, few methods capable of simultaneous analysis of a large number of analytes have been reported in pharmaceutical and environmental studies. Most of the reported analytical methods focus on one or a few analytes, the majority of which methods are performed with reversed phase chromatography. Considering the relatively simple mode of interaction (hydrophobic) between the analyte and the sorbent, one would assume that a generic method using a C18 column and a solvent gradient from aqueous to organic would work for a large number of drugs. However, the literature abounds with not only different methods for the same class of drugs but also varied methods for a single analyte. A careful look at many of these methods

would reveal that they are not very different after all. They mainly differ in the column dimensions, particle size, column packing, column manufacturer, the solvent and the gradient.

With the appeal and the availability of individual methods, a laboratory involved in analysis of several different classes of analytes, such as a clinical-biochemical laboratory in a hospital setting, tends to adopt numerous methods. The disadvantages of this approach are listed in Table 1. This can be exemplified with anti-HIV drugs as model analytes. There are numerous methods available in the literature for the analysis of different classes of anti-HIV drugs either individually or simultaneously, using different column chemistries (e.g., C18 [24], C8 [25], or Phenyl [26]). Even with the C18 column, there are many different methods that use different solvents, different pH, different ion-pairing reagents and columns from different manufacturers [27–29]. Recently, a simple method was developed [30] using a C18 column with a simple solvent system and a linear gradient (Fig. 1), which could be used to separate 16 different anti-HIV drugs simultaneously. This new method has replaced multiple assays used previously for individual drugs, simplifying the process and making it more cost effective. Similarly, 27 benzodiazepines and related drugs in human plasma have been analyzed simultaneously for detection of poisoning and screening for driving under the influence of alcohol and narcotics [13].

To increase the analytical efficiency of pharmaceutical and environmental analyses, generic separation methods aimed at simultaneous profiling of numerous compounds have to be developed. Selected reports of this type of simultaneous analysis of drugs and environmental chemicals are listed in Table 3 along with the sample-preparation and analytical methods used.

Phan-Tuan and colleagues [31] developed a generic HPLC method for high-throughput profiling of bio-fluids. They used a short monolithic column with a rapid gradient for fast profiling of a large number of samples. If a particular sample showed interesting peaks, then the same method was used with a slow gradient to identify the peaks.

In another study, a simple reversed phase separation was used by Gutteck and Rentsch [32] for baseline separation of 18 CNS active drugs in serum samples.

Garratt et al. [33] used diethylhexylamine as an ion-pairing agent for separating 119 folates, while Piraud et al. [34] tried perfluorinated carboxylic acids for separating 76 underivatized amino acids of biological interest.

A sample expected to contain a number of stereoisomers would necessitate use of a chiral column (e.g., as used by Desai and Armstrong for baseline separation of enantiomers of 25 amino-acid and peptides [35]).

For the separation of analytes with a diverse chemical functionality and polarity, complex and complementary separation modes (e.g., hydrophilic interaction chromatography [36], anion-cation exchange/hydrophilic interaction [37] or two dimensional normal-phase-reversed-phase separation) can be used [38].

3.2.2. Targeted detection—Separation methods for simultaneous analysis of hundreds of analytes from a single sample are now being developed. However, successful development and application of such methods also requires detectors capable of fast data-acquisition rates combined with high sensitivity and specificity. Targeted and non-targeted detection are the two major modes used in MS-based metabolomics.

In the targeted detection mode, the analytes to be detected are predetermined and have a specific mass filter assigned to each analyte, making it a very specific, sensitive approach. Triple-quadrupole (QqQ) instruments are very popular for targeted analysis of pharmaceuticals and

environmental chemicals due to their high sensitivity and specificity. Conventionally, they are used to quantitate one analyte or a few analytes using SIM or SRM. As SIM is not very specific and a metabolomics-type analysis of a large number of analytes requires a high level of specificity, SIM has been largely neglected for complex analysis. However, SRM offers high specificity and has been successfully used for quantitation of hundreds of analytes simultaneously. Modern QqQ instruments allow simultaneous measurement of as many as 300 analytes in a sample using SRM. We reported analysis of 141 analytes in a single sample using an SRM-based method [36]. A similar approach was used by Garratt et al. to monitor 119 folate species [33]. Fig. 2 shows the application of a targeted approach for the analysis of 17 glucocorticoids in commercial milk and eggs [14]. However, unlike conventional analysis, developing an LC-MS² method with so many SRMs is complicated and involves substantial pre-development work. The first step is to obtain fragmentation patterns for all the analytes of interest in both positive and negative modes. This data is then used to select at least two SRMs for each analyte. Selection of multiple SRMs is required in order to find the SRM that gives the highest signal-to-noise ratio in the real sample. Careful selection of SRMs is required for analysis of isobaric or isomeric compounds (e.g., glutamine and lysine with a molecular mass of 146), as these molecules often produce very similar products. Once the SRMs are selected, then the mixture of standard compounds is analyzed to obtain the retention time for each analyte necessary to develop the LC-MS² method, as shown in Fig. 3.

Good chromatographic separation and consideration of the data-acquisition rates and chromatographic peak widths are required for successful development of such a method. To obtain a reproducible peak shape, at least 10 data points per peak are required. For example, if the mass spectrometer is scanned with 0.1 sec per SRM, then a 1-min wide chromatographic peak in a segment with 60 SRMs will get 10 data points. If the peak width is less than 1 min (which is often the case for chromatography columns with shorter diameters and smaller particle sizes), the reproducibility is compromised unless the chromatographic method provides good separation. With instruments and columns now available with a particle size of less than 2 microns, peak shapes of less than 30 s can be achieved. In this case, selection of tandem mass spectrometers with a scanning speed of at least 10 ms is required. A sample normally has to be analyzed in both positive and negative polarity modes to cover a wide range of analytes. With hundreds of analytes being monitored in each mode, it is hard to obtain a chromatographic run time of less than 30 min. Hence, each sample requires at least 1 h for analysis when scanned separately in positive and negative modes. If a study involves a large number of samples, stability becomes an issue, as it often takes more than 24 h to analyze the set while samples sit in an autosampler. A combination of a faster scanning instrument, good chromatographic separation and automated positive/negative switching can significantly reduce the sample analysis time. In order to account for the MS detection as well as the extraction variability, at least one internal standard (preferably, a labeled stable isotope) per chromatographic segment (in both positive and negative ionization modes) is required.

A semi-targeted approach using a constant neutral loss scan, which is more sensitive and more specific than full-scan mode, was tried recently to detect glucuronides in human urine [39]. This approach would be suitable if the analytes belong to a certain chemical class and lose the same neutral fragments (e.g., phosphate, sulfate or glucuronide) in the MS² stage.

3.2.3. Non-targeted detection—The drawback of targeted analysis is that it is not a truly holistic approach. For a sample containing numerous analytes, it is not possible to prepare SRMs for all the molecules of interest in the sample, so the analytes for which there are no SRMs in the method cannot be detected. Also, it is often not possible to predict the composition of the sample, or the sample could have unknown analytes. In such cases, a non-targeted screening approach is commonly used in metabolomics to cover a broad range of analytes and also to detect and identify unknown or novel molecules [40,41].

In this detection mode, the mass spectrometer is scanned over a set m/z range (for example, m/z 100–1000 in both positive and negative ionization modes) to detect the majority of the molecules within the specified molecular mass range. Data obtained in full-scan mode with low-resolution instruments are generally used for comparative profiling. For example, one can compare if the samples are different in terms of analyte levels and composition. However, the identities of many analytes in the sample are not known and further analysis is required. Data acquired in full-scan mode with auto MS to MS² switching, using an accurate mass instrument (e.g., qTOF or Fourier transform MS), provide additional information on the elemental composition as well as fragmentation pattern of the analytes, including unknown molecules. The data obtained can then be subjected to a library search (such as NIST or in-house libraries) to identify the compounds [42], or used for *de novo* structure elucidation. As discussed in section 3.2.2., automated positive/negative switching in full-scan mode can significantly reduce the analysis time and increase the information content in the data. Full-scan chromatographic data has to be aligned in terms of peak-retention times, and the aligned data then can then be subjected to various statistical analyses. For more reproducible peak alignment, several retention-time markers must be added to a sample.

The unbiased detection of non-targeted analysis provides a holistic insight into the chemical nature of the sample. Such an approach is particularly important in pharmaceutical and environmental analysis as unexpected or unknown potentially toxic contaminants or drug metabolites could be detected and identified.

There are several examples of a non-targeted approach being applied in pharmaceutical and environmental analysis. Using non-targeted analysis with SPE-LC-qTOF-MS, Ibanez et al. detected six unknown compounds in environmental waters [43]. A successful application of non-targeted analysis for rapid detection of drug metabolites in pharmaceutical studies has been reported [44]. The advantage of such an approach is emphasized by the fact that it can detect previously unknown metabolites of even well-studied drugs.

The utility of the UPLC-TOF-based metabolomics approach was also shown for quality control and standardization of phytopharmaceuticals [45].

Fig. 4 demonstrates the application of non-targeted LC-MS analysis using a reversed phase monolithic capillary column for the detection of thousands of components in a plant-tissue extract.

The application of a non-targeted approach was also shown in screening 400 component libraries generated by combinatorial synthesis [46]. The information obtained is helpful in optimizing drug-synthesis procedures.

Finally, a non-targeted approach is also used in ‘General Unknown Screening (GUS)’ procedure in clinical and forensic toxicology, which involves identification of drug, poisons and intoxicating agents in samples without the prior knowledge of the analytes to be monitored [47].

There are therefore several areas where non-targeted profiling offers great advantages over a conventional highly-focused approach.

3.3. Data processing and analysis

The handling and the proper interpretation of the enormous amount of data generated by the metabolomics approach are very challenging and require the use of sophisticated statistical and bioinformatics tools [6]. LC-MS² data processing involves background-noise subtraction, peak alignment, and intensity normalization and transformation before performing statistical

analysis. Efficiency and accuracy of analysis requires a comprehensive approach, such as that used by Bijlsma et al. [48] for metabolomic analysis of 600 human samples. Fortunately, advances in metabolomics have promoted the development of several software algorithms for automated analysis of large data sets. The data-analysis approaches used for metabolomics have been reviewed [49,50] and will not be discussed in this article. Selected freely available bioinformatics tools, listed in Table 4, could easily be applied to pharmaceutical and environmental analysis.

With advances in technology and pressure for accelerated development of new drugs, drug-candidate screening has become a high-throughput process. Application of metabolomics combined with multivariate statistical data analysis can help to quickly move the drug candidates through the development pipeline. For example, Plumb et al. reported the utility of multivariate statistical analysis (PCA) of metabolomics data (Fig. 5) for efficient detection of drug metabolites in bio-fluids without prior knowledge of the drug administered [51]. Multivariate analysis can enhance the information extracted from the metabolomics data and help in rapid, correct detection of drug metabolites, including unknown metabolites [44].

Application of various highly sophisticated statistical methods to the analysis of complex analytical data sets is growing rapidly [52,53], and careful selection of the proper statistical tool needs to be based on the objectives of the analysis.

4. Challenges for metabolomics

Metabolomics, while providing tangible advantages over traditional approaches, also poses significant challenges. Metabolomics aims at quantitation, either targeted or non-targeted, of a large number of analytes in the same sample. This task has inherent difficulties because many metabolites in the sample are unknowns, and, for many known metabolites, there is no authentic standard available. Targeted analysis requires availability of all the analytes of interest, which, depending on the study, could be in the range 50–500. Obtaining the pure compounds could be difficult, as some analytes may not be commercially available while others are very costly. Targeted analysis requires substantial method development and validation work. Stock solutions, working solutions and mixtures of all the analytes must be prepared in a short period of time (2–3 days) so that all the metabolites are simultaneously available for method development, validation or quality control. If some of the analytes are unstable, fresh solutions and mixtures need to be prepared periodically.

Careful selection of optimized chromatography conditions is necessary and may be time consuming. For example, analysis of acid, base and neutral molecules, as well as polar and non-polar molecules present in the sample, requires screening of a number of chromatography solvents, columns and gradients to select the most efficient combination of factors affecting separation. Hence, a compromise in terms of chromatography performance is needed, as all the analytes may not show acceptable chromatography.

Validation of the method can be complicated as all the analytes need to be validated simultaneously and the presence of reactive analytes could affect reproducibility. Multiple internal standards are needed to represent each class of analyte.

With a non-targeted approach, it is often difficult to identify the analytes, and it requires considerable complementary analytical approaches and, again, the availability of a large number of standard compounds.

Another challenge for metabolomics is the presence of unknowns in the sample, so a complementary research effort needs to focus on structural elucidation of the unknowns. With hundreds of analytes being analyzed in a single run, it is difficult, with current technology, to

obtain faster separation times, so it is hard to achieve the high throughput required for large studies. This raises sample stability and storage issues. Manual data analysis can be complicated, and fast, proper data analysis requires novel, sophisticated computer algorithms.

Metabolomics is a promising and rapidly developing area of research. Despite current limitations, it still provides an opportunity to re-think current analytical methodology and offers researchers in pharmaceutical and environmental sciences an additional research platform that is robust and powerful.

5. Summary

High-throughput and high-information-content analysis is common in genomics, proteomics and now metabolomics. Metabolomics and pharmaceutical and environmental analysis currently share analytical techniques, which suggests a natural partnership.

A metabolomics approach applied to pharmaceutical and environmental analysis offers several advantages, such as low cost of analysis, high analytical efficiency and high information content in the data obtained. It also allows monitoring of unexpected and unpredicted metabolic changes regularly omitted by traditional analysis.

With constant technological advances in both analysis and data processing, metabolomics will provide an alternative platform to pharmaceutical and environmental analysis. In return, pharmaceutical and environmental research will add to metabolomics by providing more data on a wide range of known metabolites. As libraries grow and more metabolites are catalogued, a targeted metabolomics approach will take precedence.

Acknowledgments

The research in the authors' laboratory is financially supported by the NIH/NIGMS grant R01 GM068947-01, NIH/NIAID grant 2R01AI045774-06A2, and NSF grants MCB-0312857 and MCB-0520140. The authors thank Jim Walke for reading the manuscript critically.

References

1. Harrigan GG. IDrugs 2006;9:28. [PubMed: 16374729]
2. Walsh MC, Brennan L, Malthouse JP, Roche HM, Gibney MJ. Am J Clin Nutr 2006;84:531. [PubMed: 16960166]
3. Oresic M, Vidal-Puig A, Hanninen V. Expert Rev Mol Diagn 2006;6:575. [PubMed: 16824031]
4. Keun HC. Pharmacol Ther 2006;109:92. [PubMed: 16051371]
5. Kirchherr H, Kuhn-Velten WN. J Chromatogr, B 2006;843:100.
6. Shulaev V. Brief Bioinform 2006;7:128. [PubMed: 16772266]
7. Hollywood K, Brison DR, Goodacre R. Proteomics 2006;6:4716. [PubMed: 16888765]
8. Fiehn O. Plant Mol Biol 2002;48:155. [PubMed: 11860207]
9. Maharjan RP, Ferenci T. Anal Biochem 2003;313:145. [PubMed: 12576070]
10. Suzuki T, Nakagawa Y, Takano I, Yaguchi K, Yasuda K. Environ Sci Technol 2004;38:2389. [PubMed: 15116845]
11. Lenz EM, Bright J, Knight R, Westwood FR, Davies D, Major H, Wilson ID. Biomarkers 2005;10:173. [PubMed: 16076731]
12. Hao C, Lissemore L, Nguyen B, Kleywegt S, Yang P, Solomon K. Anal Bioanal Chem 2006;384:505. [PubMed: 16365778]
13. Kratzsch C, Tenberken O, Peters FT, Weber AA, Kraemer T, Maurer HH. J Mass Spectrom 2004;39:856. [PubMed: 15329838]

14. Cui X, Shao B, Zhao R, Yang Y, Hu J, Tu X. *Rapid Commun Mass Spectrom* 2006;20:2355. [PubMed: 16835849]
15. Aranibar N, Ott KH, Roongta V, Mueller L. *Anal Biochem* 2006;355:62. [PubMed: 16762305]
16. Castrillo JJ, Hayes A, Mohammed S, Gaskell SJ, Oliver SG. *Phytochemistry* 2003;62:929. [PubMed: 12590120]
17. Sato S, Soga T, Nishioka T, Tomita M. *Plant J* 2004;40:151. [PubMed: 15361149]
18. Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L. *Nat Biotechnol* 2000;18:1157. [PubMed: 11062433]
19. Wu L, Mashego MR, van Dam JC, Proell AM, Vinke JL, Ras C, van Winden WA, van Gulik WM, Heijnen JJ. *Anal Biochem* 2005;336:164. [PubMed: 15620880]
20. Gamache PH, Meyer DF, Granger MC, Acworth IN. *J Am Soc Mass Spectrom* 2004;15:1717. [PubMed: 15589749]
21. Dunn WB, Bailey NJ, Johnson HE. *Analyst (Cambridge, UK)* 2005;130:606.
22. Villas-Boas SG, Mas S, Akesson M, Smedsgaard J, Nielsen J. *Mass Spectrom Rev* 2005;24:613. [PubMed: 15389842]
23. Boernsen KO, Gatzek S, Imbert G. *Anal Chem* 2005;77:7255. [PubMed: 16285673]
24. Hsieh SM, Yu HY, Chang SC. *J Formos Med Assoc* 2004;103:191. [PubMed: 15124046]
25. Fan B, Stewart JT. *J Pharm Biomed Anal* 2002;28:903. [PubMed: 12039632]
26. Hirabayashi Y, Tsuchiya K, Kimura S, Oka S. *Biomed Chromatogr* 2006;20:28. [PubMed: 15971289]
27. de R, Estrela CE, Salvadori MC, Suarez-Kurtz G. *Rapid Commun Mass Spectrom* 2004;18:1147. [PubMed: 15150840]
28. Marzolini C, Telenti A, Buclin T, Biollaz J, Decosterd LA. *J Chromatogr, B* 2000;740:43.
29. Li W, Coombs RW, Collier AC, Raisys VA. *Ther Drug Monit* 1999;21:360. [PubMed: 10365654]
30. Notari S, Bocedi A, Ippolito G, Narciso P, Pucillo LP, Tossini G, Donnorso RP, Gasparrini F, Ascenzi P. *J Chromatogr, B* 2006;831:258.
31. Pham-Tuan H, Kaskavelis L, Daykin CA, Janssen HG. *J Chromatogr, B* 2003;789:283.
32. Gutteck U, Rentsch KM. *Clin Chem Lab Med* 2003;41:1571. [PubMed: 14708881]
33. Garratt LC, Otori CA, Tucker GA, Sablitzky F, Bennett MJ, Barrett DA. *Rapid Commun Mass Spectrom* 2005;19:2390. [PubMed: 16047318]
34. Piraud M, Vianey-Saban C, Petritis K, Elfakir C, Steghens JP, Bouchu D. *Rapid Commun Mass Spectrom* 2005;19:1587. [PubMed: 15915446]
35. Desai MJ, Armstrong DW. *J Mass Spectrom* 2004;39:177. [PubMed: 14991687]
36. Bajad SU, Lu W, Kimball EH, Yuan J, Peterson C, Rabinowitz JD. *J Chromatogr, A* 2006;1125:76. [PubMed: 16759663]
37. Strege MA, Stevenson S, Lawrence SM. *Anal Chem* 2000;72:4629. [PubMed: 11028621]
38. Dugo P, Favoino O, Luppino R, Dugo G, Mondello L. *Anal Chem* 2004;76:2525. [PubMed: 15117193]
39. Lutz U, Lutz RW, Lutz WK. *Anal Chem* 2006;78:4564. [PubMed: 16808466]
40. Hong S, Gronert K, Devchand PR, Moussignac RL, Serhan CN. *J Biol Chem* 2003;278:14677. [PubMed: 12590139]
41. Tohge T, Nishiyama Y, Hirai MY, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe DB, Kitayama M, Noji M, Yamazaki M, Saito K. *Plant J* 2005;42:218. [PubMed: 15807784]
42. Bobeldijk I, Vissers JP, Kearney G, Major H, Van Leerdam JA. *J Chromatogr, A* 2001;929:63. [PubMed: 11594404]
43. Ibanez M, Sancho JV, Pozo OJ, Niessen W, Hernandez F. *Rapid Commun Mass Spectrom* 2005;19:169. [PubMed: 15593265]
44. Idborg H, Edlund PO, Jacobsson SP. *Rapid Commun Mass Spectrom* 2004;18:944. [PubMed: 15116421]
45. Chan EC, Yap SL, Lau AJ, Leow PC, Toh DF, Koh HL. *Rapid Commun Mass Spectrom* 2007;21:519. [PubMed: 17238214]

46. Yates N, Wislocki D, Roberts A, Berk S, Klatt T, Shen DM, Willoughby C, Rosauer K, Chapman K, Griffin P. *Anal Chem* 2001;73:2941. [PubMed: 11467539]
47. Marquet P. *Ther Drug Monit* 2002;24:125. [PubMed: 11805733]
48. Bijlsma S, Bobeldijk I, Verheij ER, Ramaker R, Kochhar S, Macdonald IA, van Ommen B, Smilde AK. *Anal Chem* 2006;78:567. [PubMed: 16408941]
49. Dettmer K, Aronov PA, Hammock BD. *Mass Spectrom Rev* 2006;26:51. [PubMed: 16921475]
50. Arita M. *Brief Funct Genomics Proteomics* 2004;3:84.
51. Plumb RS, Stumpf CL, Granger JH, Castro-Perez J, Haselden JN, Dear GJ. *Rapid Commun Mass Spectrom* 2003;17:2632. [PubMed: 14648901]
52. Crockford DJ, Lindon JC, Cloarec O, Plumb RS, Bruce SJ, Zirah S, Rainville P, Stumpf CL, Johnson K, Holmes E, Nicholson JK. *Anal Chem* 2006;78:4398. [PubMed: 16808447]
53. Smilde AK, Jansen JJ, Hoefsloot HC, Lamers RJ, van der Greef J, Timmerman ME. *Bioinformatics* 2005;21:3043. [PubMed: 15890747]
54. Pang GF, Cao YZ, Zhang JJ, Fan CL, Liu YM, Li XM, Jia GQ, Li ZY, Shi YQ, Wu YP, Guo TT. *J Chromatogr, A* 2006;1125:1. [PubMed: 16782108]
55. Yang S, Cha J, Carlson K. *J Chromatogr, A* 2005;1097:40. [PubMed: 16298184]
56. Lopez-Roldan P, Lopez de Alda MJ, Barcelo D. *Anal Bioanal Chem* 2004;378:599. [PubMed: 13680063]
57. Rook EJ, Hillebrand MJ, Rosing H, van Ree JM, Beijnen JH. *J Chromatogr, B* 2005;824:213.
58. Petrovic M, Gros M, Barcelo D. *J Chromatogr, A* 2006;1124:68. [PubMed: 16759662]
59. Kratzsch C, Peters FT, Kraemer T, Weber AA, Maurer HH. *J Mass Spectrom* 2003;38:283. [PubMed: 12644990]
60. Maurer HH, Tenberken O, Kratzsch C, Weber AA, Peters FT. *J Chromatogr, A* 2004;1058:169. [PubMed: 15595665]
61. Broeckling CD, Reddy IR, Duran AL, Zhao X, Sumner LW. *Anal Chem* 2006;78:4334. [PubMed: 16808440]
62. Duran AL, Yang J, Wang L, Sumner LW. *Bioinformatics* 2003;19:2283. [PubMed: 14630657]
63. Katajamaa M, Oresic M. *BMC Bioinformatics* 2005;6:179. [PubMed: 16026613]
64. Spasic I, Dunn WB, Velarde G, Tseng A, Jenkins H, Hardy N, Oliver SG, Kell DB. *BMC Bioinformatics* 2006;7:281. [PubMed: 16753052]
65. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. *Anal Chem* 2006;78:779. [PubMed: 16448051]
66. Kanaya S, Kinouchi M, Abe T, Kudo Y, Yamada Y, Nishi T, Mori H, Ikemura T. *Gene* 2001;276:89. [PubMed: 11591475]

Biographies

Sunil Bajad is a Metabolomics Specialist in Vladimir Shulaev's laboratory at the Virginia Bioinformatics Institute, Blacksburg, VA, USA. He is working on the development of highly-parallel analytical methods for metabolomics and lipidomics. He holds a Ph.D. in pharmaceutical sciences with postdoctoral experience in pharmacogenomics and metabolomics. He is the author of a patent, a book chapter and several peer-reviewed research articles. His interests are development of highly-parallel LC-MS-based analytical methods and application of these methods to metabolomics and lipidomics of malarial parasite and cancer cells.

Vladimir Shulaev is Associate Professor at the Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA. He holds dual Ph.Ds. in Biological Sciences and Plant Biology, and heads the Biochemical Profiling Group, which focuses on developing methods for high-throughput metabolite profiling and application of metabolomics to systems biology, host-pathogen interactions and to study stress response in microorganisms, plants and animals.

Figure 1. A simple reversed phase method for simultaneous analysis of 16 anti-HIV drugs in human plasma (Reproduced with permission from [30]; © 2005 Elsevier B.V.).

Figure 2.

Targeted LC-MS²-based metabolomics-type approach for analysis of 17 glucocorticoids in commercial milk and eggs. The selected reaction monitoring (SRM) chromatograms shown were obtained by reversed phase ultra-performance liquid chromatography (UPLC) analysis of an egg sample spiked with standards (Reproduced with permission from [14]; © 2006 John Wiley & Sons Limited).

Figure 3.

Schematics showing comparison of conventional versus metabolomics type approach for analysis of a large number of analytes in a single liquid chromatography (LC) run using triple-quadrupole instrument. **A.** Conventional analysis of one or few analytes does not require complex LC-MS² method set-up. **B.** Metabolomics type approach with poor separation results in limitations with regard to the number of data scans per peak and the number of compounds that can be analyzed. **C.** Optimal LC-MS² method that allows simultaneous analysis of hundreds of compounds. Chromatographic run is divided into a number of time segments. Each segment can have a number of scan events, typically 1–10 and each scan event can include number of selected ion monitoring (SIM) or selected reaction monitoring (SRM) scans, typically 1–10. Peaks on the boundary of segments need to be included in both the segments. The number of SRM scans per segment is limited to about 60 in order to obtain sufficient number of data points per peak and to obtain reproducible peak shapes (with scan time of 0.1 sec, each peak will get one data point per 6 s). For details of this approach, readers are directed to [36].

Figure 4.

A non-targeted capillary LC-MS metabolomics approach used for analysis of hundreds of components in a plant-tissue extract. The profile obtained can be used for quality control and standardization of phytopharmaceuticals. The separation of the Arabidopsis tissue extract was performed on a 450×0.1-mm RP-18 capillary monolithic column using ThermoFisher LTQ mass spectrometer operated in the ESI mode with constant positive/negative switching. Data were processed with the ACD MS Manager software using CODA algorithm. A total of 730 components were detected in positive-ion mode (A) and 411 components in negative-ion mode (B).

Figure 5. Metabolomics-type approach for the analysis of drug metabolites in bio-fluids. The data shown here was obtained by applying principle component analysis to the LC-MS data obtained from urine of both control rats and rats dosed (with GSK-X). The paired numbers for each data point represent retention time (top smaller number) and mass (bottom higher number). The bold cross in the circle indicates the metabolite (retention time 6.2 min and mass of 319.13) responsible for the clustering of control and dosed groups (Reproduced with permission from [51]; © 2003, John Wiley & Sons Limited).

Table 1

Metabolomics approach compared to conventional analysis of multiple analytes

Conventional approach	Metabolomics approach
Multiple columns	Only one or two columns for multiple analytes
Different solvents	Same solvents
Different gradients	Same gradient
Time spent per sample for analysis of multiple analytes is high	Time spent per sample for analysis of multiple analytes is low
Low cost of analysis	High cost of analysis
Unknown or unexpected molecules may not be detected	Unknown or unexpected molecules may be detected
Slower data-analysis rate if multiple analytes	Fast, automated data analysis with computer algorithms
Low information content in the data	High information content in the data

Table 2

Differences in sample collection and processing for conventional vs. metabolomics analysis

Step	Conventional	Metabolomics
Collection	Does not normally require a rapid and immediate quenching step	Might require rapid and immediate quenching
Storage	Long-term storage is possible in most cases	Long-term storage may not be possible
Extraction focus	One or few analytes with similar physicochemical properties	A large number of analytes with diverse physico-chemical properties
Extraction objective	Removing majority of sample components	Retaining majority of sample components
Extraction efficiency	Can be very high and optimal	May not be very high and optimal for all the analytes
Solid-phase extraction	Possible	May not be possible
Matrix effects	Can be greatly minimized	Difficult to avoid

Table 3

Representative publications with simultaneous analysis of pharmaceuticals or environmental chemicals

Matrix	Number of components analyzed	Extraction/clean-up technique	Analytical technique	Ref.
Animal tissues	69 pesticides	Gel permeation with S-X3 column	Reversed phase chromatography with triple-quadrupole MS	[54]
Influent and effluent domestic wastewater	11 tetracycline and sulfonamide antibiotics	SPE with Oasis HLB	Reversed phase chromatography with ion-trap MS	[55]
River, ground and drinking water	21 endocrine disruptors	On-line SPE with PLRP-s cartridge	Reversed phase chromatography with single-quadrupole MS	[56]
Human plasma	12 pharmaceuticals and metabolites	SPE with Oasis MCX	Reversed phase chromatography with triple- quadrupole MS	[57]
Commercial milk and eggs	17 glucocorticoids	Liquid-liquid extraction followed by sequential SPE with Oasis HLB followed by Sep-Pak silica followed by aminopropyl cartridge	Reversed phase UPLC with triple-quadrupole MS	[14]
Surface water	27 pharmaceuticals	SPE with HLB cartridge	Reversed phase chromatography with triple-quadrupole MS	[12]
Urban wastewater	29 pharmaceuticals	SPE with Oasis HLB	Reversed phase UPLC with qTOF MS	[58]
Human serum	48 CNS active drugs	Protein precipitation by acetonitrile/ methanol	Reversed phase chromatography with triple-quadrupole MS	[5]
Human plasma	16 anti-HIV drugs	On-line SPE with Oasis HLB cartridge	Reversed phase chromatography with UV	[30]
Human plasma	18 neuroleptics and metabolites	SPE with HCX mixed-mode cartridge	Reversed phase chromatography with ion-trap MS	[59]
Human plasma	27 benzodiazepines and related drugs	Liquid-liquid extraction	Reversed phase chromatography with ion-trap MS	[13]
Human plasma	22 beta-blockers	SPE with HCX mixed mode cartridge	Reversed phase chromatography with ion-trap MS	[60]

Table 4
Selected open-access tools for complex LC-MS data analysis

Software name	Reference	Web address
MET-IDEA	[61]	http://noble.org/
MSFACTs	[62]	http://noble.org/
MZmine	[63]	http://mzmine.sourceforge.net/
MeMo	[64]	http://dbkgroup.org/memo/
XCMS	[65]	http://metlin.scripps.edu/download/
BL-SOM	[66]	http://prime.psc.riken.jp/?action=blsom_index