Mass spectrometry–based metabolomics, analysis of metabolite-protein interactions, and imaging

Do Y up Lee1, Benjamin P. Bowen1, and Trent R. Northen1,2
1Department of GTL Bioenergy and Structural Biology, Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA and 2Joint BioEnergy Institute (JBEI), Emeryville, CA, USA

Keywords: mass spectrometry; metabolomics; functional genomics; enzyme assays

Our understanding of biology has been greatly improved through recent developments in mass spectrometry, which is providing detailed information on protein and metabolite composition as well as protein-metabolite interactions. The high sensitivity and resolution of mass spectrometry achieved with liquid or gas chromatography allows for detection and quantification of hundreds to thousands of molecules in a single measurement. Where homogenization-based sample preparation and extraction methods result in a loss of spatial information, mass spectrometry imaging technologies provide the in situ distribution profiles of metabolites and proteins within tissues. Mass spectrometry–based analysis of metabolite abundance, protein-metabolite interactions, and spatial distribution of compounds facilitates the high-throughput screening of biochemical reactions, the reconstruction of metabolic networks, biomarker discovery, determination of tissue compositions, and functional annotation of both proteins and metabolites.

Sample preparation
Sample preparation is critical to metabolite analysis. While methods vary depending on the experimental goal, they are typically divided into the following steps: (i) quenching to halt metabolism, (ii) cell harvesting through medium removal (in the case of microorganism and animal cell cultures), (iii) cell lysis, and (iv) metabolite extraction. Commonly, cellular metabolism is quenched through the addition of cold methanol to a culture broth or flash-filling of tissues in liquid nitrogen. This procedure can be skipped by utilizing a filtration method that enables fast separation of cells from the medium. For metabolite extraction, a wide range of solvent systems and temperatures are used (from boiling ethanol to cold mixtures of organic solvents); however, methods of choice include cold methanol and mixtures of methanol and chloroform. Many metabolites have insufficient vapor pressure, even at high temperature (i.e., sugars, amino acids), and must be chemically modified (derivatized typically via silylation or alkylation) to increase their volatility and stability. However, chemical modification is not well-suited for thermally labile molecules or for chemicals lacking derivatizable groups (e.g., amino, hydroxyl, or carboxyl groups). Lee and Fiehn’s article offers detailed discussion and protocols for GC/MS metabolite analysis.

The development of ESI has been a breakthrough for MS analysis of intact biomolecules. The technique allows desorption and ionization of a wide range of molecules directly from the liquid phase. Therefore, it can be directly interfaced with either LC or CE. ESI is based on the formation and drying of charged liquid droplets. Fine droplets are formed through a charged nebulizer needle. As solvent evaporates, the charges are concen-
trated on the surface. When surface charge repulsion exceeds surface tension, droplets break up into smaller droplets. This consecutive concentration process continues until gas phase ions are ultimately formed.

MALDI is another widely used ionization approach for biomolecules. MALDI is performed using a large excess of compounds (matrix) that absorb light and sublime into the gas phase. These matrices efficiently assist desorption and ionization of a broad range of analytes for subsequent mass spectrometric analysis. However, the matrix itself ionizes and is detected, which greatly complicates analysis for ions with mass-to-charge ratio ($m/z$) <500 Da. It should be noted that the composition of the MALDI matrix can be modified to reduce these matrix background effects (15–17), and there have been recent innovations in matrix-free technologies. For example, nanostructure-initiator MS (NIMS) is a new technology for metabolite analysis. Here, metabolites are adsorbed onto a vacuum-compatible initiator liquid-coated nanostructured surface (18). This surface, when irradiated with a laser, vaporizes the initiator causing the desorption and ionization of analytes. It should be noted that the adsorption of nonpolar metabolites to a hydrophobic NIMS surfaces has been found to reduce signal suppression in complex biological samples (18).

**Tandem MS**

Fragmentation experiments [tandem MS (MS/MS or MSn)] are important to biomolecule analysis and enable the comparison of experimental fragmentation patterns with authentic standards and spectral databases to confirm molecular structure. The existence of metabolite spectral databases, many of which include fragmentation patterns and retention time indices, facilitates identification using GC/MS and LC/MS/MS [i.e., Golm library, Fiehn library, National Institute of Standards and Technology (NIST), Merlin, and MassBank] (19). In cases where metabolites are not in such databases, a knowledge of the precursor (unfragmented) ion’s exact mass, the exact mass of the fragment ions, and their corresponding isotopic distributions is useful for metabolite identification (20).

Fragmentation information can be acquired through various combinations of mass analyzers to isolate and fragment target ions and to analyze/detect the resulting fragments. Some of the most commonly used tandem mass spectrometers include quadrupole time-of-flight (Q-TOF), triple quadrupole (QQQ), ion trap, and Fourier transform ion cyclotron resonance (FT-ICR). Q-TOF MS instruments have high resolution, mass accuracy, and scan rate, enabling MS profiling and MS/MS analysis within a single experiment. In this case, the first quadrupole isolates ions of interest, which are then fragmented in a collision cell and analyzed by a TOF mass analyzer. QQQ mass analyzers possess three quadrupoles in a row. The first isolates, the second fragments, and the third analyzes the ions. This makes targeted MS analysis of precursor ions and MS/MS analysis of metabolites by product ion monitoring possible. This instrument can also be used for selected reaction monitoring (SRM), which provides excellent sensitivity and quantitation. Linear ion trap mass analyzers are commonly used for analysis of low abundance compounds due to their high trapping capacity as well as in MSn experiments to gain additional structural information through repeated rounds of fragmentation (21,22). While linear ion traps have low mass accuracy, FT-ICR-MS offers the highest level of resolution (>1,000,000), excellent mass accuracy (typically ~1 ppm), and good dynamic range (~5000) (23). The related orbitrap mass spectrometer has been another major breakthrough. It provides comparable mass accuracy (2–5 ppm), mass resolution (150,000), and $m/z$ range of 6000 to FT-ICR-MS, but has a considerably faster scan rate. The orbitrap and FT-ICR mass spectrometers are usually operated as a hybrid instrument with a linear ion trap (24).

**Targeted and global metabolomics**

Metabolomic approaches are often divided into targeted and untargeted. As the name suggests, targeted methods (25) are designed to detect and often quantify specific metabolites of interest within a sample. This approach has the advantage
of maximizing the specificity and the sensitivity of MS methods. Consequently, the targeted analyses utilize analytical standards to define suitable GC or LC methods, determine metabolite fragmentation patterns, and construct calibration curves for absolute quantification.

Clinical diagnostics is an early example of targeted metabolomics (26), where methods were developed to measure amino acids, drug metabolites, and specific endogenous compounds. Other targeted approaches include measurement of enzymatic activities in vitro (27) and simultaneous monitoring of multiple glycosylhydrolases and glycosyltransferases (28).

Targeted methods have also been used for pharmaceutical development, the validation of enzymatic activities for putative enzymes (29,30), and for the identification of specific substrates for putative enzymes (28,31).

In contrast, untargeted global metabolite profiling aims to maximize coverage of metabolites, often compromising the sensitivity and specificity for any particular metabolite. These metabolomic approaches involve less up-front method development when compared with targeted approaches, but require much more data analysis. The overall metabolomic workflow for untargeted LC/MS is summarized in Figure 1, and the reader is referred to detailed protocols to assist in implementation (32).

In a metabolomics experiment, sample preparation, chromatographic conditions (33), and MS ionization are all optimized to maximize the diversity of metabolites detected (34). Interpretation of the hundreds or thousands of resulting ions is challenging due to a large number of unknowns, and their identification is further complicated by the many experimental artifacts (i.e., adducts, neutral losses, isotopes). Hence, analysis depends extensively on computational tools, statistical methods, and metabolite databases.

The first step of untargeted metabolomic data analysis is to define features: the combination of the exact m/z and the corresponding LC retention time (RT). These m/z × RT dimensions subsequently are used as initial metabolite identifiers that can be quantitatively compared with important features for further analysis. There is a wide range of algorithms for identification and comparison of features, as recently reviewed by De Vos et al. (32). The most widely used algorithms are XCMS (35), msInspect (36), and mzMine (37). Identified features are quantitatively compared using univariate and multivariate statistics to select the most important features for final identification (e.g., principle component analysis (PCA) (38), partial least-squares discriminant analysis (PLS-DA) (39), and self-organizing networks (40)).

Subsequent feature identification relies primarily on exact mass searches against metabolite databases such as KEGG (41), Metlin (19), Golm (42), and HMDB (43). Database searching often results in multiple identifications for a particular ion due to insufficient mass accuracy and/or degenerate empirical formulas for a given exact mass (i.e., isomers). In these cases, identification requires analytical standards to further define the molecule’s retention time and fragmentation pattern. Unfortunately, only a relatively small subset of metabolites is commercially available. Thus, identification of unknown ions often requires either preparative-scale HPLC to enrich for NMR studies or chemical synthesis to compare with unknowns using MS/MS.

Several recent advances in the application of LC/MS and GC/MS have simplified the identification procedure: stable isotope-labeled metabolites are often used to trace pathways within metabolic networks (44). The resulting shift in m/z for the heavier metabolites is used to specify the labeling position (45), to discriminate between alternative pathways, and to estimate metabolic flux (46). In a recent example, metabolic transformations of the antiretroviral drug indinavir in vivo have been studied using knowledge-based predictions coupled with subsequent LC/MS/MS analysis (47). This study identified 18 metabolites of indinavir after incubation of the drug with human hepatic postmitochondrial preparations.

In GC/MS studies, classification and annotation of unidentified features has been improved through extraction of substructural information to predict target classes (48). Decision tree (DT)–based prediction algorithm makes use of mass spectral fragmentation and retention index (RI) information to enable sensitive and precise detection of compound substructures. It has also been shown that higher intensities of molecular and total ions (≥10-fold) are achieved by modulating a beam-steering voltage of the ion source at 70 eV EI of GC-TOF (e.g., alkanes, fatty acid methyl esters, and trimethylsilylated metabolites) (49). In this application, the accurate masses and isotopic data allowed more precise calculations of elemental compositions and facilitated metabolite identification when combined with fragmentation patterns from EI data.

### MS analysis of metabolite-protein interaction

MS is being used for annotation of putative enzyme functions, suggesting potential metabolic reactions, and validating the existence of metabolic pathways using in vitro enzymatic assays. Such studies can be broken down into three groups, as shown in Figure 2.

#### Solution-based in vitro binding or enzyme assays

This is a metabolomics profiling experiment that examines the loss or appearance of metabolites following enzymatic reactions. A limitation of this approach is that these metabolites often need to be detected within a complex sample matrix (i.e., in a cell lysate). Recently, this approach

---

**Figure 2. MS analysis of protein-metabolite interaction.** (A) ABPP utilizes chemical probes and beads to selectively profile actively bound proteins through specific crosslinking reactions and proteomic analysis of tagged peptides. (B) Protein arrays are used to bind ligands that can subsequently be released for MS-based identification. (C) Surface-immobilized metabolites can bind proteins (proteomic analysis) or detect specific enzymatic transformations of the immobilized substrate.
Immobilized protein-based assays

Here, one or more metabolites are immobilized and used as bait to identify their interacting protein partner(s) either by physically isolating the protein or by using affinity chromatography. This approach was used to identify 11 interaction partners with the new metabolite N\textsubscript{4}-(N-acetylamino)propyl spermidine by immobilizing the metabolite on agarose beads, binding the protein(s), and elution for subsequent proteomic identification (57). Activity and selectivity estimates can be made using this tag-free ABPP strategy (click chemistry) (53,54).

Immobilized metabolite binding/enzyme assays

Here, one or more metabolites are immobilized and used as bait to identify their interacting protein partner(s) either by physically isolating the protein or by using affinity chromatography. This approach was used to establish 11 interaction partners with the new metabolite N\textsubscript{4}-(N-acetylamino)propyl spermidine by immobilizing the metabolite on agarose beads, binding the protein(s), and elution for subsequent proteomic identification (57).

Another recent example is the Nimzyme assay, which uses fluoruous phase interactions for soft immobilization of the enzyme substrates onto a NIMS (58). This approach allows direct characterization of crude cell extracts, since the immobilized substrates and products are retained on the surface. It has been shown to be suitable for detection of both addition and cleavage reactions (sialyltransferase and galactosidase). In a related approach, alkanethiol monolayers are used for substrate immobilization on metal surfaces, and laser desorption ionization MS is used to analyze reaction products following enzyme treatments (59).

MS-based metabolite imaging

A limitation of LC/MS, CE/MS, or GC/MS methods is the loss of spatial information that results upon metabolite extraction from homogenized samples. Metabolomic imaging technologies, therefore, can be an important alternative and provide information on the spatial distribution of metabolites within tissues (Figure 3). MALDI imaging is the most widely used (60) MS-based tissue imaging approach. MALDI matrix is typically applied to the sample (i.e., tissue) either by spotting or spraying, and images are generated by raster scanning the laser over the sample, providing a mass spectrum at each x,y coordinate. Composite images are constructed by mapping the distribution and abundance of ions within the sample.

Another widely used approach for imaging is secondary ion MS (SIMS) (61). In SIMS, the sample is sputtered with ions (i.e., gold, gallium, bismuth) to generate secondary ions. The resultant ions are analyzed typically using time-of-flight mass analyzers (TOF-SIMS). This matrix-free technique does not have the background matrix ions and also has the advantage that ion beams can be focused to ~100 nm. However, the energetic SIMS ionization process results in extensive molecular fragmentation that significantly complicates metabolite identification and data interpretation.

Such limitations in imaging analysis are now being addressed using new soft techniques, such as desorption electrospray ionization (DESI). DESI is based on scanning the surface with the electrospray cone (62). NIMS is another matrix-free approach that uses a scanning laser system to generate ions from the liquid-filled nanostructured surface (18).

Conclusion

The application of MS to metabolite analysis is providing new insights into the biochemical functions and the cellular physiology of living organisms. These approaches are becoming widely used for biomarker discovery and systematic investigation of metabolic dynamics. Targeted approaches are used to monitor and often quantify specific metabolites of interest with high sensitivity and selectivity, whereas untargeted metabolomic profiling is allowing detection of unexpected metabolites. Specific reactions can be probed using activity-based metabolite/protein profiling and MS-based enzyme assays, and now developments in MS-based imaging are enabling the determination of metabolite spatial distribution within tissues.

Acknowledgements

We gratefully acknowledge support from the Department of Energy (DE-AC02-05CH11231), the National Institutes of Health (NIH), the National Cancer Institute (NCI), Bay Area Breast Cancer SPORE (P50 CA 58207), and the California Breast Cancer Research Program (151B-0063). We also thank Deron Herr at Scripps Research Institute images. This paper is subject to the NIH Public Access Policy.

Competing interests

The authors declare no competing interests.

References


domonas reinhardtii. Plant Methods 4:7.


16. McCombie, G. and B. Knochelmuss. 2004. Small-molecule MALDI using the matrix suppression effect to reduce or eliminate matrix background inter-


formatics 8:105.


25. Bennett, B.D., J. Yuan, E.H. Kimball, and J.D. Rabionowitz. 2008. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-


