Anal. Chem. **2009,** *81,* 6546–6551

^{13}C Isotope-Labeled Metabolomes Allowing for Improved Compound Annotation and Relative Quantification in Liquid Chromatography-Mass Spectrometry-based Metabolomic Research

Patrick Giavalisco,* Karin Köhl, Jan Hummel, Bettina Seiwert, and Lothar Willmitzer

Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

Metabolomics is rapidly becoming an integral part of many life science studies ranging from disease diagnostics to systems biology. However, a number of problems such as the discrimination of biological from non-biological signals, efficient compound annotation, and reliable quantification are still not satisfactorily solved in untargeted LC-MS-based metabolomics research. Extending our previous work on direct infusion-based metabolomics, we here describe a ^{13}C isotope labeling strategy in combination with an Ultra Performance Liquid Chromatography Fourier Transform Ion Cyclotron Resonance Mass Spectrometry-based approach (UPLC-FTICR MS) which provides a technological platform offering solutions to a number of the above-mentioned problems. We further demonstrate that the use of a fully labeled metabolome is not only beneficial for high end mass spectrometers, such as that used in this study but also provides a considerable improvement to every other mass spectrometry-based metabolomic platform.

Metabolomics aims at the identification and quantification of all metabolites present in a given biological system.[1] This can be achieved using either mass spectrometry- (MS) or nuclear magnetic resonance (NMR) spectroscopy-based techniques.[2−5] Both technologies have advantages and are therefore complementary.[3,4] However, because of its high sensitivity and the capability to analyze highly complex samples, MS-based approaches have started to take the lead in metabolomic research.[4]

Two main types of study can be distinguished in MS-based metabolomics, namely, targeted and untargeted approaches.[6−9] The major difference between these approaches is that the targeted approach will, in the best case, consider only those compounds which are described for the biological system of interest. The untargeted approach instead considers every measured signal a possible metabolite and therefore aims to interpret it. While the restricted targeted analysis, which depends on mass spectral and chromatographic comparison of authentic reference substances to measured sample specific features,[7] does not suffer from annotation or quantification problems, untargeted approaches are still limited by a number of problems including accurate compound annotation and quantification.

We have recently described a comparative $^{12}C/^{13}C$ isotope labeling strategy for *Arabidopsis thaliana* plants, allowing fast screening/annotation of metabolites, using direct infusion-based FTICR MS.[10] On the basis of this approach, we were able to annotate 1,024 non-redundant elemental compositions. However, as a result of the chromatography-free, direct infusion-based sample delivery into the mass spectrometer, the formerly presented method still suffers from a number of limitations. Structural isomers, sharing the same elemental composition and therefore mass, could not be discriminated from each other. Fragments and multimers, generated during the ionization process cannot be distinguished from or associated to their parent ions, since the origin of these masses cannot be traced back without an accurate, compound-specific chromatographic retention time. Last but not least, a feasible option for the accurate relative quantification of the measured compounds is still lacking. To overcome these problems we extended our initial approach by incorporating a reproducible and sensitive chromatographic system, making the whole process also accessible to lower resolution mass spectrometers.

EXPERIMENTAL SECTION

Chemicals. All chemicals used were purchased from Sigma-Aldrich (Steinheim, Germany) with the highest purity grade available, while the solvents were purchased from BioSolve (Valkenswaard, Netherlands), unless otherwise stated.

Plant Growth and Metabolite Extraction. The *A. thaliana* plants and the analyzed leaf tissue was grown and extracted as described in the previous publication.[10]

UPLC-NanoMate-FTICR MS Measurement. UPLC separation was performed using a Waters Acquity UPLC system (Waters,

* To whom correspondence should be addressed. E-mail: giavalisco@mpimp-golm.mpg.de. Phone: 0049-331-5678623. Fax: 0049-331-5678236.

(1) Fiehn, O. *Plant Mol. Biol.* **2002**, *48*, 155–171.
(2) Ward, J. L.; Baker, J. M.; Beale, M. H. *FEBS J.* **2007**, *274*, 1126–1131.
(3) Pan, Z.; Raftery, D. *Anal. Bioanal. Chem.* **2007**, *387*, 525–527.
(4) Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrom. Rev.* **2007**, *26*, 51–78.
(5) Hollywood, K.; Brison, D. R.; Goodacre, R. *Proteomics* **2006**, *6*, 4716–4723.
(6) Sawada, Y.; Akiyama, K.; Sakata, A.; Kuwahara, A.; Otsuki, H.; Sakurai, T.; Saito, K.; Hirai, M. Y. *Plant Cell Physiol.* **2008**.
(7) Lu, W.; Bennett, B. D.; Rabinowitz, J. D. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2008**, *871*, 236–242.
(8) Hanhineva, K.; Rogachev, I.; Kokko, H.; Mintz-Oron, S.; Venger, I.; Karenlampi, S.; Aharoni, A. *Phytochemistry* **2008**, *69*, 2463–2481.
(9) De Vos, R. C.; Moco, S.; Lommen, A.; Keurentjes, J. J.; Bino, R. J.; Hall, R. D. *Nat. Protoc.* **2007**, *2*, 778–791.
(10) Giavalisco, P.; Hummel, J.; Lisec, J.; Inostroza, A. C.; Catchpole, G.; Willmitzer, L. *Anal. Chem.* **2008**, *80*, 9417–9425.

10.1021/ac900979e CCC: $40.75
Published on Web 07/09/2009

Mildford, MA, U.S.A.), using a HSS T3 C_{18} reverse phase column (100 × 2.1 mm i.d. 1.8 μm particle size, Waters) at a temperature of 40 °C. The mobile phases consisted of 0.1% formic acid in water (Solvent A) and 0.1% formic acid in acetonitrile (Solvent B). The flow rate of the mobile phase was 400 μL/min, and 2 μL samples were loaded per injection. Gradient conditions were: 0−1 min hold at 1% B, 1−13 min linear gradient 1%−35% B, 13−14.5 min linear gradient from 35%−70% B, 14.5−15.5 min linear gradient 70%−99% B, 15.5−17 min hold at 99% B, 17−17.5 min linear gradient to 1% B and 17.5−20 min hold at 1% B.

The UPLC was connected to the FTICR via a TriVersa NanoMate (Advion, Ithaca, NY, U.S.A.). The UPLC flow rate, which was 400 μL/min, was split 1:1000 with a T-Valve (Advion). One tenth of a percent was directly loaded to the FTICR MS, while 99.9% were delivered to waste.

400 nL/min of sample were infused into the MS via a nanospray Chip (Type A, Advion) by applying a voltage of 1.8 kV in the positive and 1.9 kV in the negative ionization mode. Spray sensing was used between min 1 and 17 of the UPLC gradient.

The mass spectra were acquired using the LTQ FTICR-Ultra mass spectrometer (Thermo-Fisher, Bremen, Germany). The spectra were recorded using full scan mode, covering a mass range from m/z 100−1300. The resolution was set to 50,000, and the maximum loading time for the ICR cell was set to 500 ms. The transfer capillary temperature was set to 200 °C, and the MS spectra were recorded from min 1 to 17 of the UPLC gradient.

Relative Quantification of ^{12}C and ^{13}C Samples. For the relative quantification of the differentially mixed ^{12}C/^{13}C samples 7 aliquots of ^{12}C-derived metabolite extracts were mixed 1:1, 1:2, 1:5, 1:10, 2:1, 5:1, 10:1 with ^{13}C metabolite extract. The samples were measured using the UPLC-FTICR MS in positive ionization mode. The spectra and resulting peaks were manually extracted and evaluated using the Xcalibur software (Version 2.06, Thermo Fischer). For peak evaluation the peak heights of 20 selected ^{12}C/^{13}C peak pairs were extracted from every mixture, and their ratios were calculated. In a next step the 1:1 ratio of each selected peak pair, which was regarded as the reference, was divided by the ratios of the same peak pair, derived from the other mixtures. These inter-sample ratios were then used to calculate the reproducibility and the accuracy of the quantitative analysis.

Data Analysis. Molecular masses, retention time, and associated peak intensities for the three replicates of each sample group (^{12}C positive, ^{12}C negative, ^{13}C positive, and ^{13}C negative) were extracted from the raw files using the SIEVE software (Version 1.1, Thermo-Fisher). The mass and retention time lists were used for searches against the ChemSpider database,[11] employing the in-house developed database search tool GoBioSpace (Hummel et al., unpublished). This tool was realized using a Microsoft SQL Server 2005 as the relational database backend for storing chemical sum formulae with appropriate source tagged annotations (names, synonyms, cross references, etc). Algorithms for formula parsing and isotopic correct mass calculations were implemented as user-defined types using the Common Language Runtime (CLR). Net framework, the C# programming language, and Microsoft Visual Studio 2005. The search criteria, which can be restricted to a mass error of between 0.1−100 ppm, were set to 2 ppm, and elemental compositions restricted to contain only C, H, N, O, P, or S.

The ^{12}C and the ^{13}C data sets were analyzed individually, and the result files, including the database annotations of each mass, associated elemental compositions, retention time, m/z value, compound ID, and possible substance names were exported as text files. Data visualization was performed using Microsoft Excel (Excel 2007, Microsoft, Redmont, WA, U.S.A.). The content of these files was sorted and filtered either directly in the GoBioSpace search tool or by using Microsoft Access (Access 2007, Microsoft). The sorting of the data included the matching of ^{12}C and ^{13}C elemental compositions within the same ionization mode and the retention time alignment of matched elemental compositions. All other spectra manipulations and peak extractions were performed using the Xcalibur software (Version 2.06, Thermo Fischer).

RESULTS AND DISCUSSION

In this technical note we describe an significantly improved analytical platform using ^{12}C and ^{13}C labeled metabolomes[10,12] in combination with ultra performance liquid chromatography (UPLC)[13−15] and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS).[16] The concept and the identification strategy of the approach are depicted in Figure 1a. From two batches of independently grown plants, either supplied with $^{12}CO_2$ or $^{13}CO_2$ as sole carbon source, metabolites were extracted and analyzed independently by UPLC FTICR MS. As shown in Figure 1b the chromatograms derived from the independent ^{12}C- and ^{13}C- samples display almost identical peak patterns. Only the mass spectra, underlying the various chromatographic peaks, indicate that the two samples are derived from two different isotopically labeled regimes, enabling the extraction of information concerning the different compounds as follows:

(1) If a compound is of biological origin, a mass shift between the co-eluting ^{12}C and the ^{13}C peak will be observed (Figure 1b peak at 7.27 min). If no mass shift is detectable (Figure 1b peak at 8.23 min), this compound cannot be assigned to be of biological origin.

(2) The number of carbon atoms of the respective compound can be derived by subtracting the integer mass of the ^{12}C peak from the integer mass of the co-eluting ^{13}C peak. In the example of the peak at 7.27 min in Figure 1b, the ^{13}C peak has a mass of m/z 774.3335, whereas the co-eluting ^{12}C peak has a mass of m/z 741.2238, indicating that this compound contains 33 carbon atoms.

This information presents a major advance, compared to non-isotope supported metabolomics, as it allows not only discrimina-

(11) Williams, A. J. *Curr. Opin. Drug Discovery Dev.* **2008**, *11*, 393–404.

(12) Hegeman, A. D.; Schulte, C. F.; Cui, Q.; Lewis, I. A.; Huttlin, E. L.; Eghbalnia, H.; Harms, A. C.; Ulrich, E. L.; Markley, J. L.; Sussman, M. R. *Anal. Chem.* **2007**, *79*, 6912–6921.

(13) Castro-Perez, J.; Plumb, R.; Granger, J. H.; Beattie, I.; Joncour, K.; Wright, A. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 843–848.

(14) Guy, P. A.; Tavazzi, I.; Bruce, S. J.; Ramadan, Z.; Kochhar, S. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2008**, *871*, 253–260.

(15) Yonekura-Sakakibara, K.; Tohge, T.; Matsuda, F.; Nakabayashi, R.; Takayama, H.; Niida, R.; Watanabe-Takahashi, A.; Inoue, E.; Saito, K. *Plant Cell* **2008**, *20*, 2160–2176.

(16) Brown, S. C.; Kruppa, G.; Dasseux, J. L. *Mass Spectrom. Rev.* **2005**, *24*, 223–231.

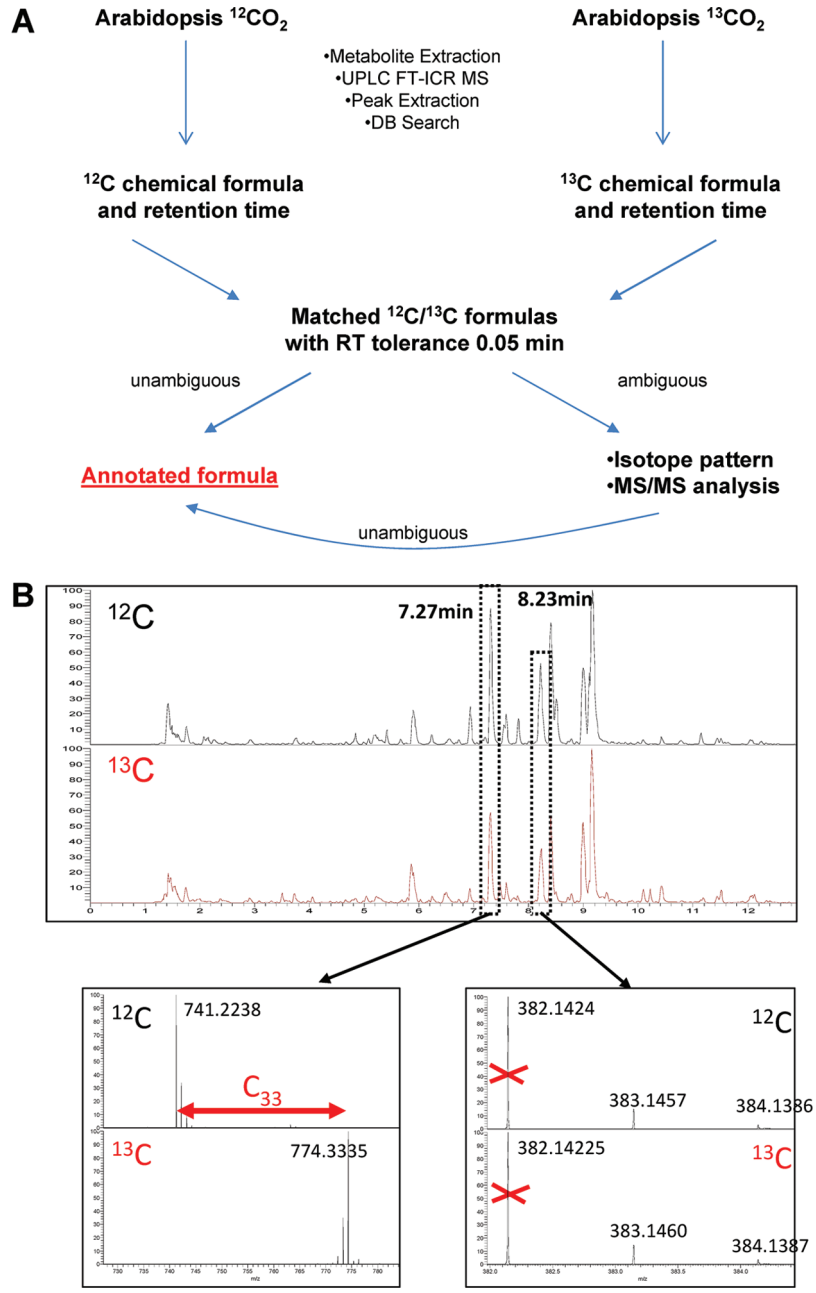

Figure 1. (A) Overview of the general strategy for the untargeted isotope labeling-based analysis of metabolomes. (B) Positive ion mode spectra of *A. thaliana* leaf metabolites. The upper chromatogram represents the total ion chromatogram (TIC) of the ^{12}C (black) and ^{13}C (red) sample. The lower left spectra show the ^{12}C and ^{13}C mass spectra at the apex of the chromatographic peak at 7.27 min, while the right-hand spectrum shows the ^{12}C and ^{13}C mass spectra of the peak at 8.31 min. The presence/absence of a mass shift between the displayed masses in the upper (^{12}C) and lower (^{13}C) row of each mass spectrum indicates (1) the biological /non-biological origin of the compound, and (2) in case a mass shift is detectable, the number of carbon atoms of the measured compound.

tion between peaks from the biological sample and contaminations but also supports the unambiguous annotation of the elemental composition of this compound. The annotation of the ^{12}C peak of m/z 741.2238 (Figure 1b), when compared with the ~14 Million entries containing ChemSpider DB,[11] resulted in three different elemental compositions (1 ppm search tolerance and limiting the search to $[M+H^+]$ adducts). Increasing the error tolerance to a mass accuracy of 10 ppm increases the number of matching compositions already to 16 (Supporting Information, Figure 1a). However, introducing information on the number of carbon atoms eliminates all ambiguities up to an error tolerance of 10 ppm, thus leading to a single chemical sum formula (Supporting Information, Figure 1b). This is an important result since it demonstrates that our approach is applicable not only to ultra high resolution/high mass accuracy mass spectrometers such as the FTICR MS or the Orbitrap MS[17−19] but also to instruments with considerably lower mass accuracy.[4,20]

After demonstrating the principle of the applicability of our approach to chromatography-based analysis, we then applied the method to the same sample set, that was used with the previously published direct infusion-based FTICR MS analysis.[10] Three independent biological replicates of $^{12}CO_2$ and $^{13}CO_2$ grown plants

Table 1a. Peak Statistics of the Different Samples Used for the $^{12}C/^{13}C$ Chemical Formula Annotation[a]

sample	no. of peaks	DB hits	M+H	M+Na	M+NH_4	M−H	M+FA−H
^{12}C negative	4189	24323	×	×	×	11264	13059
^{12}C positive	2658	25442	8529	9041	7872	×	×
^{13}C negative	3536	19136	×	×	×	8889	10247
^{13}C positive	2386	21969	7151	8112	6706	×	×

[a] **No. of peaks** indicates how many chromatographic peaks were extracted from a total of three aligned replicates. **DB hits** summarizes the total number of hits, which were obtained if the no. of peaks was searched against the ChemSpider[11] database. **M+H** stands for a molecule ionizing with a proton as an adduct., **M+Na** indicates a molecule being ionized with sodium as an adduct. **M+NH_4**, indicates molecules ionizing with ammonia as an adduct. All three adducts form cation ions. **M−H** and **M+Fa-H** are negative ions, formed by the loss of a proton. M+Fa-H additionally, next to the loss of a proton contains a formic acid adduct (derived from the UPLC solvent). The numbers below each of these columns gives the number of DB hits for each of these adducts.

Table 1b. Summary of Elemental Composition Annotations after Having Matched the $^{12}C/^{13}C$ Chemical Formulas[a]

ionization mode	overlap chem. formula	overlap after RT correction	RT groups	differential formulas
positive	5688	1832	365	1377
negative	4281	1299	357	1090

[a] **Ionization mode** indicates the measurement type the samples were derived from. **Overlap chem. formula** summarizes all possible combinations of matching chemical formula pairs, while **overlap after RT correction**, displays the number of chemical formulas that are matching and elute within a maximally allowed retention time window (tolerance 3 s). **RT groups** display the number of distinct retention time groups, while **differential formulas** display the total number of chemical formulas extracted from the chromatogram.

were analyzed in both positive and negative ionization mode. The measured chromatograms resulted in up to 4,200 extractable chromatographic peaks per sample and ionization mode (Table 1a). The accurate masses associated with these extracted peaks were used for comprehensive ^{12}C or ^{13}C specific database searches against the ChemSpider database.[11] These searches resulted in more than 20,000 ambiguous elemental compositions per sample/ionization mode (Table 1b).

The availability of two differentially isotope labeled plants permitted now the application of a number of simple yet powerful filtering options which allow us not only to exclude the nonbiological formulas and reduce therefore the number of ambiguous elemental compositions but also to distinguish structural isomers and annotate the fragment and parent ions (Figure 1a). The first step of this filtering was achieved by matching the formulas derived from the different ^{12}C and the ^{13}C database searches (^{12}C positive to ^{13}C positive and ^{12}C negative to ^{13}C negative), which led to 4,281 negative mode and 5,688 positive mode peak pairs (Table 1b).

These peak pairs with matching elemental composition were obtained without yet haven taken into account that identical compounds, in contrast to structural isomers, must co-elute and therefore possess highly similar retention times (Figure 1a). Allowing a maximal retention time difference of three seconds (0.05 min) between pairs of elemental compositions appearing in the ^{13}C respective ^{12}C chromatograms, aligns the matching isomers and therefore provides a bona fide result list of 1,299 negative- and 1,832 positive mode-derived elemental compositions derived from the same compound (Table 1b). Interestingly, in accordance to previously published data from human serum metabolomics,[21] we could also observe a strong independence (less than 10% overlap) of the chemical formulas derived from the positive and the negative ionization mode measurements, underlining the strong complementarity of these two measurement modes (data not shown).

That said, this list of chemical formulas still overestimates the number of truly distinct biological compounds present in *A. thaliana* leaves. Many compounds were annotated with more than one ionization adduct and/or produced different breakdown- or multimerization products. This still amplifies the number of "true" database hits and $^{12}C/^{13}C$ peak- and RT pairs. To get a better estimate of the number of truly different compounds, we used the concept of retention time grouping. Essentially mass traces displaying identical retention times and peak shapes can be regarded as signals derived from a single precursor. Figure 2 graphically illustrates, based on two representative examples, how the distinction between co-eluting, associated, and non-associated masses is determined. Obviously all the masses present in the spectrum of compound I and the upper four masses in the spectrum of compound II are connected, based on their identical chromatographic behavior (Figure 3c). In contrary the masses m/z 358.1146 and m/z 919.2556 of compound II are derived from different compounds because they have distinguishable chromatographic peaks with different retention times (Figure 3c).

When taken together, the number of distinguishable chromatographic retention time groups containing different numbers of associated masses (Figure 2), results in 357 different negative- and 365 different positive ionization mode-derived peaks detectable in the chromatograms from *A. thaliana*, constituting a total number of 643 different retention time pairs of unique elemental compositions (Table 1b). Interestingly the total number of different elemental compositions annotated to the 643 UPLC-FTICR MS-derived retention time groups was with a number of 2,304 more than twice as high as the 1,024 direct infusion-derived elemental compositions[10] (Table 1b). This underlines the strength of the chromatography-based approach, namely, not only gaining precision in the mapping of chemical formulas, by reducing them to a distinct number of retention time groups, but also gaining

(17) Scheltema, R. A.; Kamleh, A.; Wildridge, D.; Ebikeme, C.; Watson, D. G.; Barrett, M. P.; Jansen, R. C.; Breitling, R. *Proteomics* **2008**, *8*, 4647–4656.
(18) Makarov, A.; Denisov, E.; Lange, O.; Horning, S. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 977–982.
(19) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R. *J. Mass Spectrom.* **2005**, *40*, 430–443.
(20) Werner, E.; Heilier, J. F.; Ducruix, C.; Ezan, E.; Junot, C.; Tabet, J. C. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2008**, *871*, 143–163.
(21) Nordstrom, A.; O'Maille, G.; Qin, C.; Siuzdak, G. *Anal. Chem.* **2006**, *78*, 3289–3295.

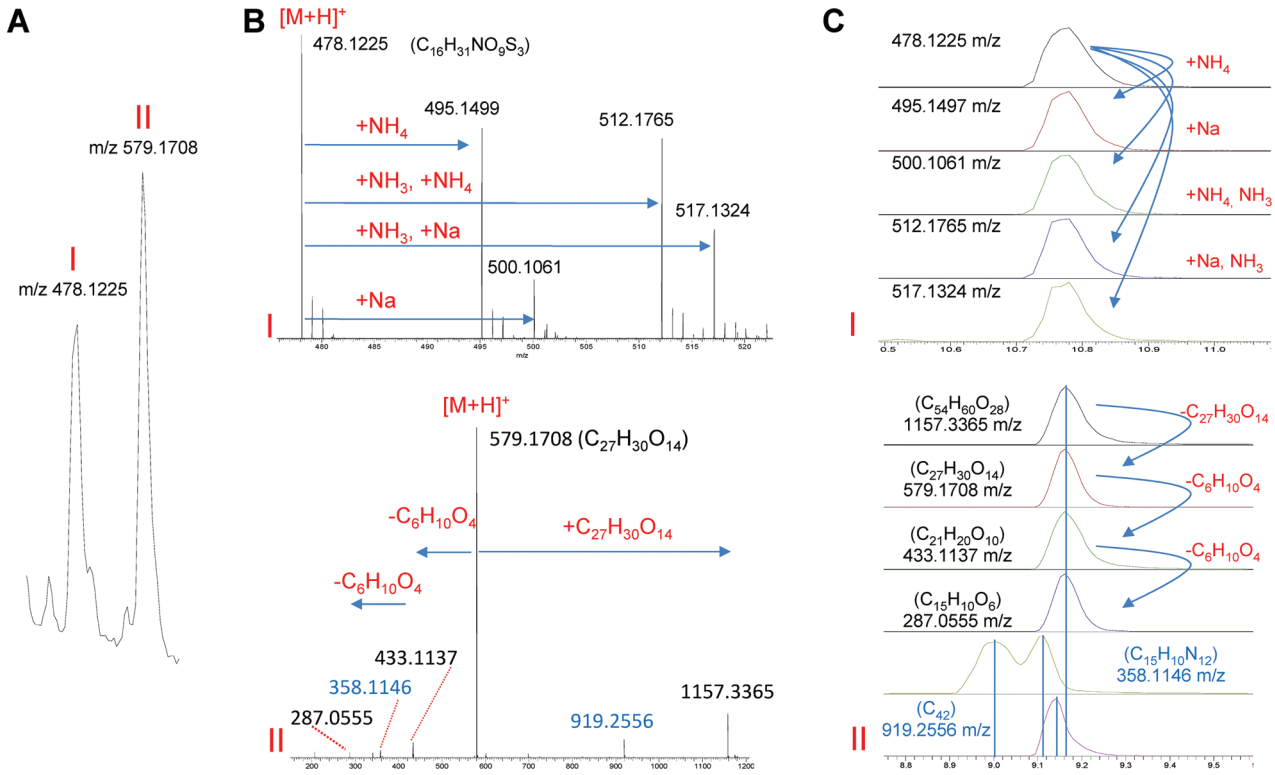

Figure 2. (A) Base peak chromatogram of two neighboring peaks from a positive ion mode spectrum of *A. thaliana*. Peak I is the chromatographic peak of a compound with the elemental composition $C_{16}H_{31}NO_9S_3$ and a *m/z* of 478.1225. Peak II corresponds to a compound with a mass of *m/z* 579.1708 and the chemical formula $C_{27}H_{30}O_{14}$. (B) Mass spectrum from the apex of peak I and II as depicted in A. Mass shifts are indicated by blue arrows and their possible cause is depicted as a chemical formula in red above each arrow. (C) Extracted ion chromatograms (EIC) visualizing the chromatographic behavior of masses present in the apex of the peaks I and II (see B). Each row represents the EIC of a mass (indicated at the left-hand side of each line), while the blue arrows at the right-hand side of each row indicate the molecular cause of the mass shift between the displayed masses.

sensitivity because of reduced ion suppression in the UPLC-based measurement.[22,23]

To obtain a more quantitative overview concerning the level of ambiguity within the chemical sum formula annotation using either the ^{12}C masses alone or with a carbon number restriction, we researched the masses of the top 98 $^{12}C/^{13}C$ peak/retention time pairs, with different mass tolerances, against the ChemSpider[11] database. As expected the number of ambiguous formulas decreases significantly with increasing mass accuracy.[24] As a consequence we found only 1 unambiguous elemental composition if no carbon number restriction was applied to a 10 ppm database search. The number of unambiguous database hits increases slightly to 6 for a 5 ppm search while a 2.5 ppm database search resulted in 15 unambiguous hits (Table 2). Only the database search using an error tolerance of 1 ppm led to a significant increased number of unambiguous elemental compositions per mass, namely, 53 (Table 2).

Restricting the carbon number now for all database searches, based on the $^{12}C/^{13}C$-derived carbon number, resulted in a significant increase in unambiguous elemental composition annotations. Even the 10 ppm search now provided a number of 71 masses returning a single elemental composition from the database search, while the 5 ppm, 2.5 ppm, and 1 ppm database searches, resulted in 84, 88, and 96 unambiguous database hits, respectively (Table 2). This data demonstrates the extreme usefulness of the isotope labeling-derived carbon number information for the efficient mass to elemental composition annotation.

A closer look at the 10 ppm data reveals that 80% of the 27 ambiguously assigned masses from the database search resulted in only two possible elemental compositions. Interestingly most of them could be manually annotated based on the fact that they either contained sulfur or not. Sulfur containing compounds can be easily discriminated from non-sulfur containing compounds based on their characteristic isotope distribution with a pronounced M+2 peak.[25]

A further significant advantage of our UPLC-FTICR MS-based isotope labeling approach, compared to the previously published direct infusion-based method, is that it allows for accurate relative quantification. Akin to concepts which have been described in proteomic research,[26−28] our approach allows for the relative quantitative analysis of a large number of different ^{12}C samples. This relative quantification can be achieved by simply spiking each sample, which usually will be a ^{12}C grown sample, with a fixed amount of a ^{13}C-derived metabolite extract. The relative

(22) Annesley, T. M. *Clin. Chem.* **2003**, *49*, 1041–1044.
(23) Muller, C.; Schafer, P.; Stortzel, M.; Vogt, S.; Weinmann, W. *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2002**, *773*, 47–52.
(24) Clauser, K. R.; Baker, P.; Burlingame, A. L. *Anal. Chem.* **1999**, *71*, 2871–2882.
(25) Shi, S. D.; Hendrickson, C. L.; Marshall, A. G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11532–11537.
(26) Harsha, H. C.; Molina, H.; Pandey, A. *Nat. Protoc.* **2008**, *3*, 505–516.
(27) Ong, S. E.; Mann, M. *Nat. Chem. Biol.* **2005**, *1*, 252–262.

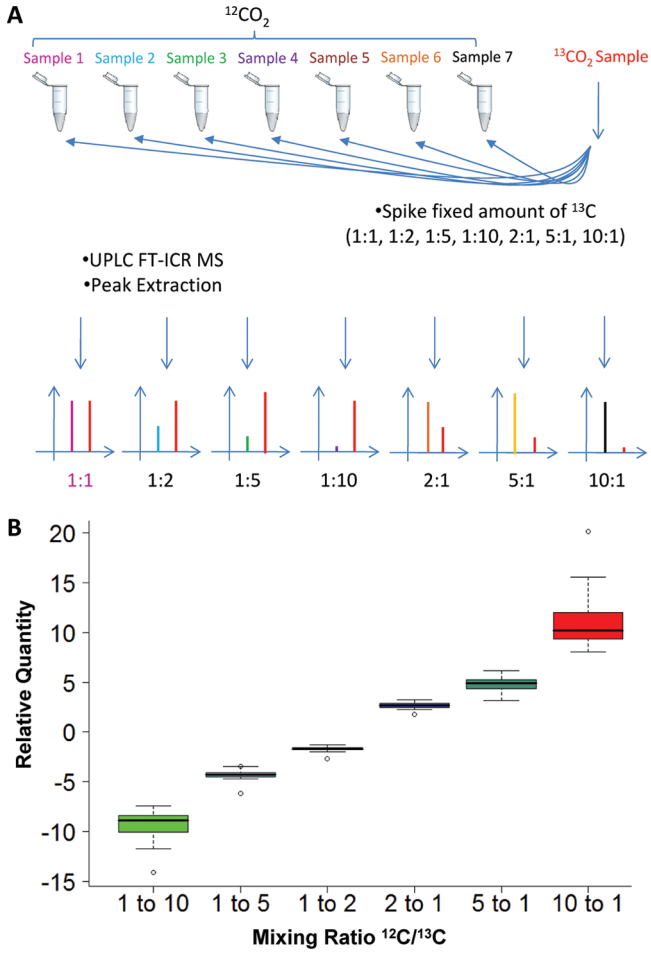

Figure 3. (A) Overview of the general strategy for the relative quantification of different ^{12}C samples using identical ^{13}C labeled metabolomes as a complex internal standard mixture. The ratios in the figure, like, e.g., 1:1, 1:2, represent the ^{12}C:^{13}C ratio of the mixtures. (B) Box plots of the measured $^{12}C/^{13}C$ ratios of 20 representative metabolites observed in the dilution series (expected ratios are given on the *x*-axis). Values for the 1:10, 1:5, and 1:2 samples are negative and reciprocal to improve the presentability. The data of this diagram is derived from Supporting Information, Table 2.

Table 2. Distribution of Ambiguous versus Unambiguous Chemical Formula Annotations for the ^{12}C Masses of the 98 Most Abundant $^{12}C/^{13}C$ Peak Pairs[a]

	10 ppm	5 ppm	2.5 ppm	1 ppm
no carbon number restriction	1%	6%	15%	53%
with carbon number restriction	71%	84%	88%	96%

[a] The displayed results show the percentage of unambiguous annotated masses, if searched against the ChemSpider database.[11] The first line represents the results if all chemical formulas matching if the masses are searched with column wise displayed ppm tolerance, while the second row shows the percentage if the masses derived from the database search are corrected by the carbon number derived from the $^{12}C/^{13}C$ mass shift of the corresponding peak.

amount of each peak can then be quantified between different samples by simply comparing the $^{12}C/^{13}C$ ratios between the different samples for the peaks of interest. According to this the spiked ^{13}C extract can be regarded as an extremely complex isotopically labeled internal standard.

To test and demonstrate the feasibility of our approach we designed a proof of concept experiment which is illustrated in Figure 3a. In this experiment we mixed a ^{12}C *A. thaliana* metabolite extract in different ratios (1:1, 1:2, 1:5, 1:10, 2:1, 5:1, and 10:1) with a ^{13}C-labeled *A. thaliana* metabolite extract. For these mixtures we recorded the UPLC-FTICR MS spectra and analyzed the quantitative behavior of 20 manually extracted $^{12}C/^{13}C$ peak pairs (Supporting Information, Table 1). To validate the quantitative accuracy of the measurements we used the $^{12}C/^{13}C$ 1:1 ratio of each of the 20 selected peak pairs and divided it by the ratio of the corresponding $^{12}C/^{13}C$ peak pair from the other mixtures. These obtained ratios are summarized in the Supporting Information, Table 2. As can be seen from the Supporting Information, Table 2, and its graphical illustration in Figure 3b, the results of the peak ratios of the analyzed $^{12}C/^{13}C$ peaks provide an excellent quantitative measure with an average relative standard deviation of 15% for the 6 different $^{12}C/^{13}C$ dilutions (Supporting Information, Table 2). This data indicates that ^{13}C isotope spiking does not only provide good precision and reproducibility but also allows subtle differences like in the 1:2 or 2:1 mixture to be accurately and reproducible detected and quantified.

CONCLUSION

We here provide an improved strategy for truly untargeted qualitative and quantitative metabolomics. The described update is based on the utilization of differentially ^{13}C isotope labeled metabolomes in combination with highly reproducible and sensitive UPLC-based chromatographic separation. The presented approach leads to a significantly improved elemental composition annotation (as a first step toward compound identification), discrimination of biological from non-biological compounds, and comprehensive, accurate relative quantification for differential analysis of an unlimited number of ^{12}C samples. Even though the approach presented here was developed on an ultra high resolution FTICR MS, it would improve metabolomic research significantly even if less sophisticated mass spectrometers were used, thus making untargeted metabolomics accessible to a broad scientific community.

ACKNOWLEDGMENT

We would like to acknowledge Dr. Leonard Krall, Dr. Gareth Catchpole, and Dr. Alisdair Fernie for critical comments on the manuscript. Aenne Eckardt is kindly acknowledged for the excellent help with sample preparation and UPLC-MS measurements. Carsten Richter of Gaswechsel-Mess-Systeme GmbH developed the "BioBox light" isotope labelling system and provided excellent support.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review May 6, 2009. Accepted June 11, 2009.

AC900979E

(28) Ong, S. E.; Mann, M. *Nat. Protoc.* **2006**, *1*, 2650–2660.