Method for Determining Molar Concentrations of Metabolites in Complex Solutions from Two-Dimensional ¹H−¹³C NMR Spectra

Ian A. Lewis, Seth C. Schommer, Brendan Hodis, Kate A. Robb, Marco Tonelli, William M. Westler, Michael R. Sussman, and John L. Markley*

Department of Biochemistry, University of Wisconsin—Madison, 433 Babcock Drive, Madison, Wisconsin 53706

One-dimensional (1D) ¹H nuclear magnetic resonance (NMR) spectroscopy is used extensively for high-throughput analysis of metabolites in biological fluids and tissue extracts. Typically, such spectra are treated as multivariate statistical objects rather than as collections of quantifiable metabolites. We report here a two-dimensional (2D) ¹H−¹³C NMR strategy (fast metabolite quantification, FMQ, by NMR) for identifying and quantifying the ∼40 most abundant metabolites in biological samples. To validate this technique, we prepared mixtures of synthetic compounds and extracts from *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, and *Medicago sativa*. We show that accurate (technical error 2.7%) molar concentrations can be determined in 12 min using our quantitative 2D ¹H−¹³C NMR strategy. In contrast, traditional 1D ¹H NMR analysis resulted in 16.2% technical error under nearly ideal conditions. We propose FMQ by NMR as a practical alternative to 1D ¹H NMR for metabolomics studies in which 50-mg (extract dry weight) samples can be obtained.

One-dimensional (1D) ¹H nuclear magnetic resonance (NMR) spectroscopy has been used for decades as an analytical tool for identifying small molecules and measuring their concentrations.[1,2] Traditionally, quantitative analysis by NMR has been restricted to relatively simple mixtures with minimal peak overlap. In these applications, 1D ¹H NMR is a natural choice, because its peaks scale linearly with concentration and its analytical precision is usually independent of the chemical properties of target molecules. Recently, interest has surged in using NMR for high-throughput analysis of complex biological processes at the metabolic level.[3,4] These studies, defined as "metabolomics" or "metabonomics", place an emphasis on biomarker discovery or disease classification and are typically centered on unfractionated biological fluids and tissue extracts. 1D ¹H NMR spectra of these samples typically contain hundreds of overlapping resonances (Figure 1) that make traditional NMR-based analytical practices, such as resonance assignment and accurate peak integration, a challenging prospect. As a result, sophisticated statistical tools have been developed to translate spectral data into biologically meaningful information.[4,5]

All statistical tools used for analyzing complex spectra face the same fundamental barrier: overlapped peaks do not scale in the discrete linear fashion that typifies well-isolated peaks. They scale as the sum of the total overlapped resonance. Consequently, multivariate and correlation statistics are reporters of overlapped spectral density, not concentrations of specific compounds. Although peak overlap does not interfere with the reproducibility of traditional analyses,[6] it does prevent accurate quantification.

Two approaches can be used to overcome this barrier, one mathematical and the other experimental. The mathematical approach is to fit overlapped 1D NMR spectra with modeled peaks. This approach has been successfully applied by Weljie and co-workers.[7] The experimental approach is to collect NMR spectra that disperse peaks into two or more dimensions. This allows non-overlapped peak intensities to be measured directly.

Two-dimensional NMR (Figure 1) is a well-established technique for reducing peak overlap and has been recognized for over 10 years as an excellent tool for metabolomics.[8−11] Despite this, published applications of multidimensional NMR in the metabolomics literature have been largely restricted to qualitative analyses. One reason for this is that 2D cross-peak intensities (or volumes) are influenced by a greater number of variables (e.g., uneven excitation, nonuniform relaxation, evolution times, mixing times, etc.) than are 1D ¹H NMR peaks. This nonuniform behavior makes it difficult to translate peak intensities into metabolite concentrations. A second reason is that 2D NMR spectra usually require more time to collect than 1D ¹H spectra. Long NMR acquisition times are impractical for metabolomics studies requiring the analysis of hundreds of samples.

* Corresponding author. Tel: 608-263-9349. Fax: 608-262-3759. E-mail: markley@nmrfam.wisc.edu.

(1) Pauli, G. F.; Jaki, B. U.; Lankin, D. C. *J. Nat. Prod.* **2005,** *68,* 133−149.

(2) Radda, G. K.; Seeley, P. J. *Annu. Rev. Physiol.* **1979,** *41,* 749−769.

(3) Lindon, J. C.; Nicholson, J. K.; Holmes, E.; Everett, J. R. *Concepts Magn. Reson.* **2000,** *12,* 289−320.

(4) Lindon, J. C.; Holmes, E.; Nicholson, J. K. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001,** *39,* 1−40.

(5) Holmes, E.; Cloarec, O.; Nicholson, J. K. *J. Proteome Res.* **2006,** *5,* 1313−1320.

(6) Dumas, M. E.; Maibaum, E. C.; Teague, C.; Ueshima, H.; Zhou, B.; Lindon, J. C.; Nicholson, J. K.; Stamler, J.; Elliott, P.; Chan, Q.; Holmes, E. *Anal. Chem.* **2006,** *78,* 2199−2208.

(7) Weljie, A. M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C. M. *Anal. Chem.* **2006,** *78,* 4430−4442.

(8) Fan, T. W. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996,** *28,* 161−219.

(9) Viant, M. R. *Biochem. Biophys. Res. Commun.* **2003,** *310,* 943−948.

(10) Kikuchi, J.; Shinozaki, K.; Hirayama, T. *Plant Cell Physiol.* **2004,** *45,* 1099−1104.

(11) Fan, T. W. M.; Lane, A. N.; Shenker, M.; Bartley, J. P.; Crowley, D.; Higashi, R. M. *Phytochemistry* **2001,** *57,* 209−221.
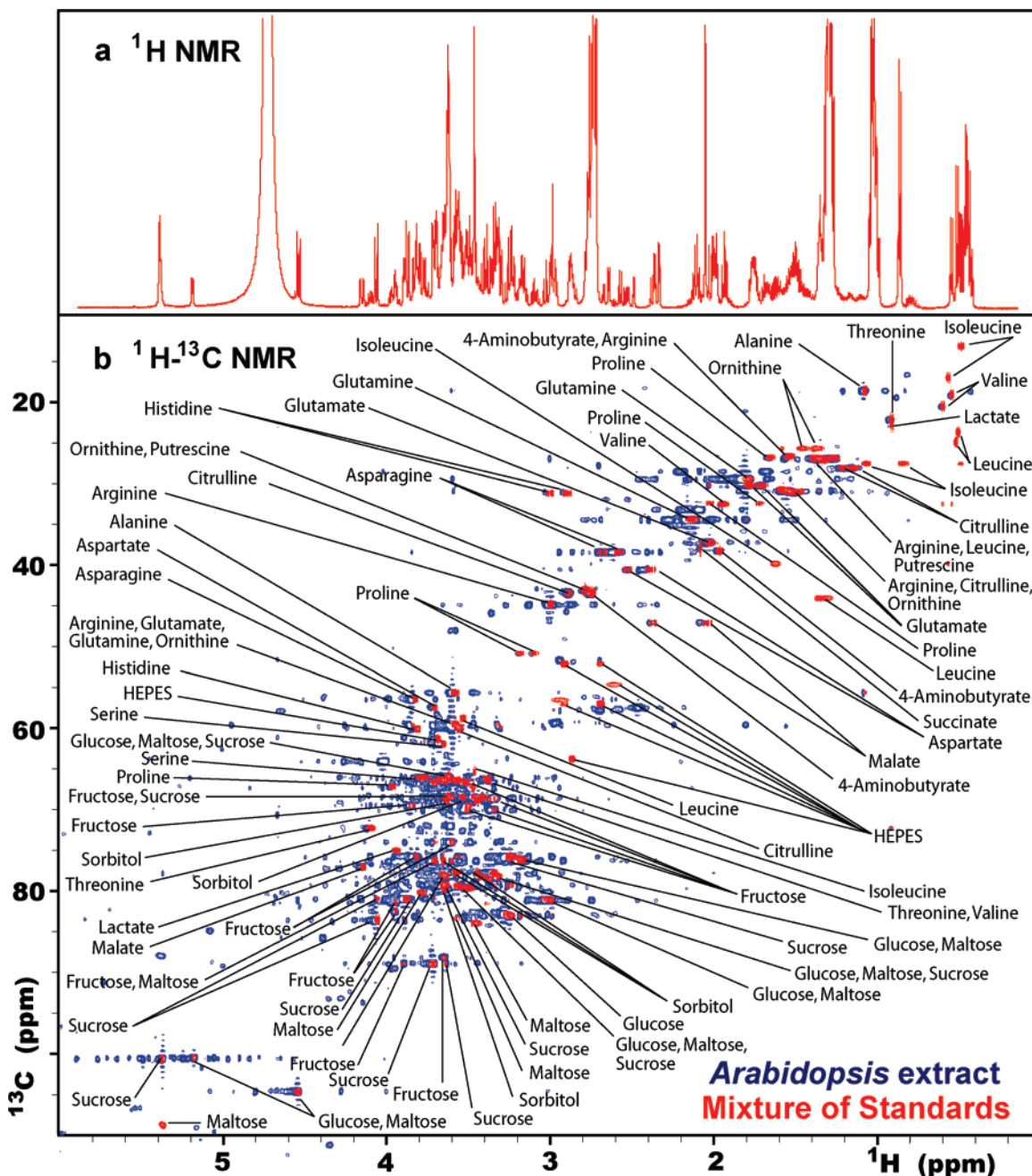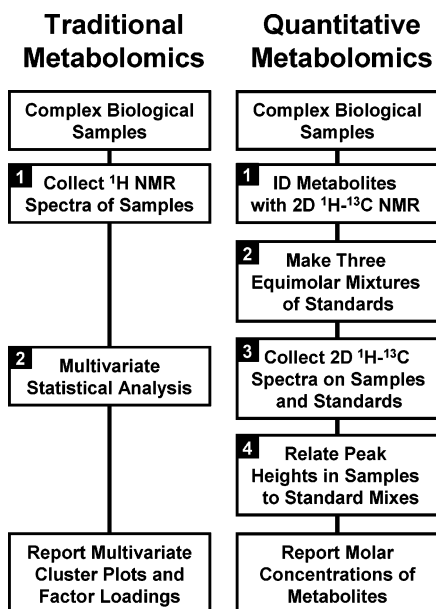
Figure 1. (a) One-dimensional ^1H NMR spectrum of an equimolar mixture of the 26 small-molecule standards listed in Supporting Information Table 1. (b) Two-dimensional $^1H-^{13}C$ HSQC NMR spectra of the same synthetic mixture (red) overlaid onto a spectrum of aqueous whole-plant extract from *A. thaliana* (blue).

In this paper, we introduce a simple experimental protocol, fast metabolite quantification (FMQ) by NMR (Figure 2), for measuring molar concentrations of metabolites in complex solutions using 2D $^1H-^{13}C$ NMR. This method is accurate and allows 2D experiments to be collected in as little as 12 min. Our method requires that all target metabolites be identified and that 2D $^1H-^{13}C$ NMR spectra of standards have been collected under comparable conditions. Metabolite identification has recently become feasible with three new public databases: the metabolomics extension of the Biological Magnetic Resonance Data Bank (BMRB, www.bmrb.wisc.edu);[12,13] the Madison Metabolomics Consortium Database (MMCD, http://mmcd.nmrfam.wisc.edu),[13] Qiu Cui, I.A.L., Adrian D. Hegeman, Mark E. Anderson, Jing Li,

Christopher F. Schulte, W.M.W, Hamid R. Eghbalnia, M.R.S, and J.L.M., *Nature Biotechnology*, in press; and the Human Metabolome Project (HMP, www.metabolomics.ca).[14] We describe how to identify metabolites using these tools and demonstrate the process by identifying and quantifying metabolites in *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, and *Medicago sativa* tissue extracts. In addition, we estimate the technical error associated with 1D and 2D NMR analyses of complex mixtures. The protocol

(12) Seavey, B. R.; Farr, E. A.; Westler, W. M.; Markley, J. L. *J. Biomol. NMR* **1991**, *1*, 217−236.
(13) Markley, J. L.; Anderson, M. E.; Cui, Q.; Eghbalnia, H. R.; Lewis, I. A.; Hegeman, A. D.; Li, J.; Schulte, C. R.; Sussman, M. R.; Westler, W. M.; Ulrich, E. L.; Zolnai, Z. *Pac. Symp. Biocomput.* **2007**, *12*, 157−168.

Traditional Metabolomics

Complex Biological Samples

1 Collect ¹H NMR Spectra of Samples

2 Multivariate Statistical Analysis

Report Multivariate Cluster Plots and Factor Loadings

Quantitative Metabolomics

Complex Biological Samples

1 ID Metabolites with 2D ¹H-¹³C NMR

2 Make Three Equimolar Mixtures of Standards

3 Collect 2D ¹H-¹³C Spectra on Samples and Standards

4 Relate Peak Heights in Samples to Standard Mixes

Report Molar Concentrations of Metabolites

Figure 2. (Left) Two-step experimental design used in traditional metabolomics, which typically reports multivariate statistics related to spectra. (Right) Our four-step quantitative metabolomics protocol, "FMQ by NMR", which yields molar concentrations for all identified metabolites.

we describe is a general quantitative procedure that can be applied to any biological system in which 50-mg (extract dry mass) samples can be obtained.

EXPERIMENTAL SECTION

Experimental Rationale. One of the main purposes of this study was to evaluate 1D ¹H NMR and 2D ¹H−¹³C NMR as quantitative tools for metabolic profiling. To do this, we identified ∼80% of the NMR-observable metabolites present in *Arabidopsis, Saccharomyces*, and *Medicago* extracts. A subset of these compounds was selected to represent an average extract, and several synthetic mixtures of pure compounds were prepared. All of the error estimates reported in this study are based on the synthetic mixtures rather than real biological extracts. This approach allowed us to measure the absolute error associated with the experimental techniques. Although our synthetic samples contained considerably fewer compounds than a real extract, the mixtures were created with biologically realistic concentrations and chemical diversity.

Plant Growth and Tissue Extraction. *A. thaliana* seeds were germinated and grown in sterile liquid cultures of MS medium.[15] Seedlings were incubated under continuous illumination on a shaker platform. After two weeks of growth, plants were harvested, washed in ddH₂O, and flash frozen in liquid nitrogen. Wild type (DS10) *S. cerevisiae* were prepared by growing a culture in YPD medium until cells had reached the stationary phase (optical

(14) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M. A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; Macinnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. *Nucleic Acids Res.* **2007**, *35*, D521−526.
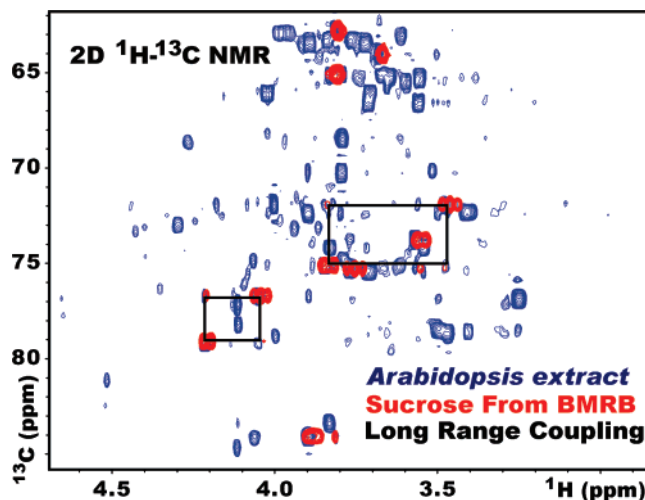(15) Marashige, T.; Skoog, F. *Physiol. Plant.* **1962**, *15*, 473−479.

density of 15). Cell suspensions were centrifuged, and the pellet was washed in 20 volumes of phosphate-buffered saline (10 mM sodium phosphate, 150 mM NaCl, pH 7.0). Washed cell suspensions were centrifuged, and the remaining cell pellet was flash frozen in liquid nitrogen. *M. sativa* sprouts were obtained from a local grocery store. Sprouts were washed in ddH₂O and flash frozen in liquid nitrogen.

Frozen *Arabidopsis* seedlings, *Medicago* sprouts, and *Saccharomyces* cells were lyophilized for 48 h and homogenized in a coffee grinder. A total of 400 mg of each dry homogenate was suspended in 16 mL of boiling ddH₂O and incubated at 100 °C in a screw-top 22-mL vial for 15 min. Extracts were microfiltered through ddH₂O-washed, 3-kDa cutoff spin filters, and the filtrate was lyophilized to a dry powder. Each 400 mg tissue sample yielded approximately 50 mg of dry metabolite extract. The lyophilized extract was suspended in NMR buffer B: D₂O, 5 mM 4-(2-hydroxylethyl)-1-piperazine ethanesulfonic acid (HEPES), 500 μM NaN₃, and 500 μM 2,2-dimethylsilapentane-5-sulfonic acid at a volume-to-weight ratio of 8.75 μL of buffer B/mg of extract. The resulting solution was titrated with DCl or NaOD as needed to achieve an observed pH of 7.400 (±0.004).

Preparation of Synthetic Samples. Thirty synthetic mixtures were prepared for the error analysis study. Twenty-four of these samples were designated as "test mixtures", and 6 samples were designated as "concentration references". Both the test mixtures and concentration references contained 26 small molecules (see Supporting Information Table 1). Twenty-five metabolite standards were selected from the larger list of molecules identified in the three biological extracts (see Quantitative Protocol, step 1). HEPES was also included in the synthetic samples as an internal concentration reference. All mixtures were prepared from weighed pure standards (Sigma-Aldrich) suspended in NMR buffer B and were titrated to an observed pH of 7.400 (±0.004). Test mixtures contained 19 metabolites with invariant concentrations (all 5 mM) and 7 metabolites with variable concentrations ranging from 5.5 to 29.1 mM. Although each test mixture had a unique metabolite profile, the samples were designed to group into six classes with biologically relevant concentrations and standard deviations (see Supporting Information Table 2). The six concentration reference samples were prepared with equimolar mixtures of the 26 metabolites. These samples contained each metabolite at 2 ($N = 2$), 5 ($N = 2$), or 10 mM ($N = 2$).

A separate set of biological concentration reference standards was prepared for estimating concentrations in the three tissue extracts. Biological reference samples had a total of 52 metabolites split between 3 groups. The groups were designed to minimize overlap between metabolites signals in 2D ¹H−¹³C NMR spectra. Biological references were prepared in the same manner as described for the synthetic concentration reference samples.

NMR Spectroscopy. All NMR spectroscopy was carried out at the National Magnetic Resonance Facility in Madison. Spectra were collected on a Varian 600-MHz spectrometer equipped with a triple-resonance (¹H, ¹³C, ¹⁵N, ²H lock) cryogenic probe and a sample changer. The probe was tuned, matched, and shimmed by hand for the first sample. All subsequent samples were collected using an automated shimming and data acquisition macro written in-house. 1D ¹H and 2D ¹H−¹³C heteronuclear single-quantum correlation (HSQC) spectra of each sample were

collected. 1D ¹H spectra were collected using 90° pulses with four acquisitions, four silent scans, an initial delay of 2 s, and an acquisition time of 2 s. Sensitivity enhanced ¹H−¹³C HSQC spectra were collected with 4 scans, 32 silent scans, an initial delay of 1 s, and an acquisition time of 0.3 s with broadband decoupling. The spectral width and number of increments were adjusted to achieve a good compromise between resolution and total acquisition time. Quantitative ¹H−¹³C HSQC spectra were collected in 128 increments using a 70 ppm spectral width in the indirect (¹³C) dimension. The carbon transmitter offset frequency was tuned to allow all aliphatic resonances to be contained within the spectrum. Our objective was to minimize spectral width while avoiding peak aliasing in the aliphatic region. Aromatic resonances and the anomeric resonances of sugars were allowed to wrap into the top of the spectral window (see Supporting Information Figure 1). Each quantitative ¹H−¹³C HSQC spectrum required 12 min to collect. With these spectrometer settings, every molecule in the synthetic mixtures yielded at least one non-overlapped cross-peak as did each of the identified metabolites in the biological extracts, with the exception of putrescine, lactate, and acetate. One high-resolution ¹H−¹³C HSQC spectrum was collected for each biological extract. These spectra were acquired with 512 increments, 16 scans, and a ¹³C spectral width of 140 ppm. High-resolution spectra were used to identify metabolites. The larger spectral width helped avoid resonance assignment errors resulting from spectral folding.

Data Processing and Statistical Analysis. All spectra were chemical-shift referenced, phased, Fourier-transformed with a shifted sine bell window function, zero-filled, and peak-picked using automated nmrPipe[16] processing scripts written in-house. Concentration calculations, regressions, and error analyses were done using automated scripts written in R, a free statistics software package (www.r-project.org). A detailed description of the calculations, annotated R scripts, and all of the raw data are available (see Data Availability). Regression analyses report the best fit linear model of calculated concentrations ($N = 168$) versus the known concentration of each metabolite. Error analyses report the mean absolute difference between calculated metabolite concentrations ($N = 168$) and known concentrations.

Protocol for FMQ by NMR. *Step 1: Identification of Metabolites in Cell Extracts.* High-resolution ¹H−¹³C HSQC data were collected from each biological extract, and nmrPipe[16] was used to process and peak-pick each spectrum (see NMR Spectroscopy). A list of possible metabolite matches was obtained by cross-referencing observed ¹H−¹³C chemical shifts (NMR peak locations) with shifts in the BMRB and MMCD databases. All matches were checked visually by overlaying spectra of pure standards from the BMRB onto the high-resolution ¹H-¹³C HSQC spectrum of each extract. Sparky (freely available from www.cgl.ucsf.edu/home/sparky) was used to prepare the overlaid spectra. All of the assigned metabolites with concentrations greater than ∼5 mM were further validated on the basis of long-range ¹H−¹³C couplings observed in the 2D spectra (Figure 3). A total of 52 metabolites met our assignment criteria, and 26 of these metabolites were included in the synthetic mixtures used in the error analysis study (see Supporting Information Table 1).

(16) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.

Figure 3. Two-dimensional ¹H−¹³C HSQC NMR spectrum of sucrose from the BMRB (red) overlaid onto an aqueous whole-plant extract from *A. thaliana* (blue). Black boxes indicate long-range proton carbon couplings used to validate the assignment.
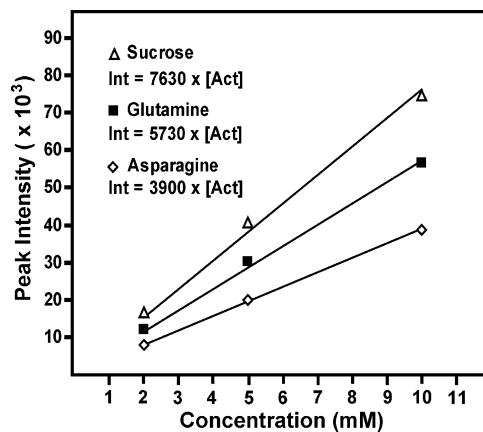
Step 2: Concentration Reference Samples. Six equimolar mixtures of the identified metabolites (2 mM $N = 2$, 5 mM $N = 2$, and 10 mM $N = 2$) were prepared as concentration reference samples (see Preparation of Synthetic Samples).

Step 3: Data Collection. All test samples and concentration reference samples were run as a block under identical acquisition parameters (see NMR Spectroscopy). The sample block was run twice to produce two technical replicates for each sample. To minimize experimental bias, a random number generator (www.random.org) was used to determine the sample order. 1D ¹H and 2D ¹H−¹³C HSQC spectra were collected sequentially on each sample with custom macros written in-house.

Step 4: Calculation of Concentrations. For the error analysis study, two well-dispersed peaks in the ¹H−¹³C HSQC spectra were selected for each of the seven target metabolites. A standard curve was constructed for each metabolite by regressing absolute peak intensities from the concentration reference samples with their known concentrations (Figure 4). Standard curves were averaged across the technical replicates ($N = 4$, two block replicates, and two sample replicates), and the resulting regression coefficients were used to predict metabolite concentrations in the test samples. Concentration estimates were also averaged across technical replicates ($N = 4$, two block replicates, and two peaks were selected from each molecule) to produce a final predicted concentration for each metabolite. Identical procedures were used to estimate concentrations from both 1D and 2D NMR data. Proton chemical shifts used in the 1D ¹H analysis were those identified from the positions of 2D ¹H−¹³C HSQC cross-peaks. These shifts were hand-verified to ensure that the correct resonance was selected. Some minor adjustments to the HSQC-based chemical shifts (±0.005 ppm) were necessary because of the higher resolution of the 1D ¹H spectra. Chemical shift translation was done with custom scripts written in R.

Step 4 (Alternative): Normalized Calculation of Concentrations. All of the samples used in this study contained 5 mM HEPES as an internal standard. An alternative strategy used for calculating concentrations was to normalize all observed intensities to the average signal from the two dispersed HEPES peaks. Standard

Figure 4. Two-dimensional $^1H-^{13}C$ HSQC NMR peak intensities for three metabolites in the concentration reference samples plotted as a function of known concentration. The concentrations of metabolites in the test samples were calculated from the best-fit regression lines of the concentration reference samples.
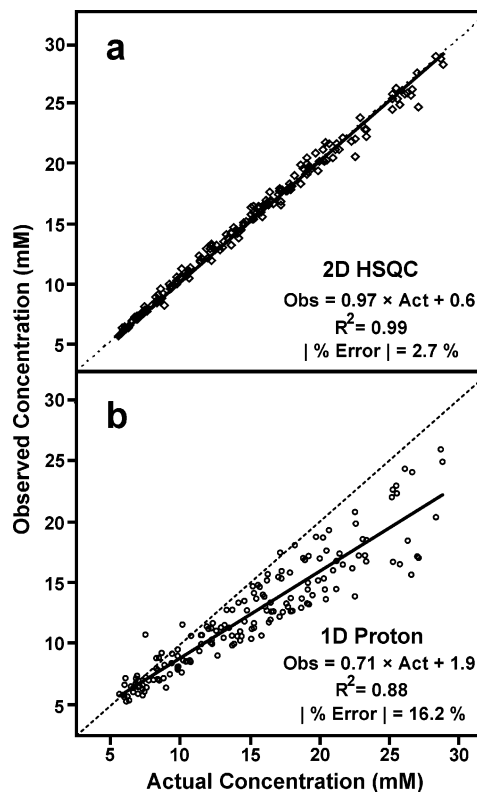

Figure 5. (a) Concentration estimates ($N = 168$) based on two-dimensional $^1H-^{13}C$ HSQC and (b) one-dimensional ^1H NMR data. Estimates are plotted as a function of the known concentration of metabolites in synthetic mixtures. Dotted lines indicate the ideal regression (slope 1), and the solid lines indicate the best-fit regression.

curves were then constructed with the HEPES normalized intensities, and concentrations were predicted in the same manner as described above. This method corrects for any changes in the NMR sensitivity between experiments and allows standards to be collected at a different time from the test samples. Metabolite concentration estimates for the biological extracts were carried out in this manner, and all well-dispersed peaks were used in the calculations. Chemical shifts and raw intensities of peaks used for quantification are provided as text files in the Supporting Information.

Data Availability. A detailed description of the metabolite identification process, assignment criteria, chemical shifts of assigned metabolites, a metabolite assignment tutorial, processed NMR data, and R functions used for automatic peak assignment, quantification, and regression analysis are available as downloadable files from the NMRFAM website at http://www.nmrfam.wisc.edu/Software/FMQ_data/.

RESULTS

Regression and Error Analyses. A total of 186 NMR-based concentration estimates were made from NMR spectra of the 24 synthetic mixtures. Each estimate was regressed against the known concentration of the metabolite (Figure 5). Parallel regressions of 2D $^1H-^{13}C$ HSQC and 1D ^1H NMR measurements indicated that the 1D estimates were considerably noisier than the 2D estimates ($r^2 = 0.88$ versus $r^2 = 0.99$). In addition, regression slopes indicated that the 2D $^1H-^{13}C$ HSQC estimates were precise (slope 0.97; theoretical slope 1), whereas the 1D ^1H-based data underestimated concentration (slope 0.71; theoretical slope 1). Error estimates calculated from the divergence of observed from actual concentrations indicated that the HSQC-based method averaged 2.7% error with a maximum of 10.3%, while the 1D ^1H analysis averaged 16.2% error with a maximum of 44.5% (Supporting Information Figure 2). This error translates to an average (root-mean-square) accuracy of 0.6 mM for the 2D and 3.5 mM accuracy for 1D estimates.

Quantification of Biological Extracts. Fifty metabolites (excluding HEPES and MES) were identified in high-resolution spectra of the three biological extracts. Forty of these were

observed in *Arabidopsis*, 37 were seen in alfalfa sprouts, and 41 were present in yeast extracts. These metabolites represented, respectively, 81, 81, and 86% of the observed peaks in the *Arabidopsis*, alfalfa, and yeast extracts (Supporting Information Table 2). Of the 50 observed metabolites, 41 (82%) could be quantified in automated fashion from 12-min 2D $^1H-^{13}C$ data acquisitions. Observed concentrations, as measured in the NMR tube, ranged from 230 mM to as little as 40 μM (Supporting Information Table 3). Metabolite concentrations below 1 mM should be treated with caution because they are below the limit of our measured error. Half of the metabolites seen in high-resolution $^1H-^{13}C$ NMR had concentrations lower than 4 mM, and metabolites observed in the three tissues were distributed proportionally to the inverse of concentration (Figure 6).

Factors Influencing Quantification. It should be noted that most metabolite concentrations were predicted well beyond the range of the standard intensity versus concentration curves. In an earlier trial of this experiment, we found that higher standard concentrations (>10 mM for each metabolite) influenced the sensitivity of the probe and produced nonlinear calibration curves. When included in an analysis, the nonlinear standards caused a systematic underestimate of concentration and greater technical error (data not shown). It should also be noted that the standards data must be collected at multiple concentrations for accurate quantification. Multidimensional peak intensities are influenced by a range of variables (uneven excitation, multiple relaxation pathways, etc.) that change the scaling factors for individual peaks from different metabolites (Figure 4). Our data show that calculat-
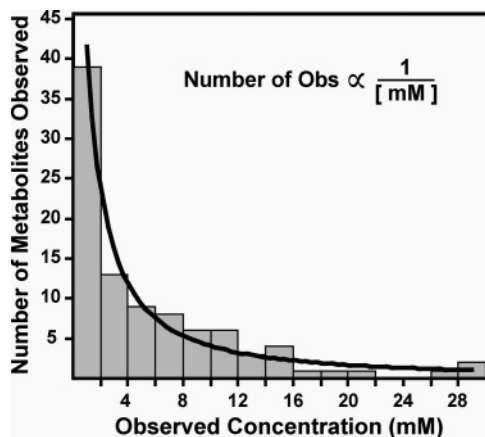
Figure 6. Number of observed metabolites in *Arabidopsis*, *Medicago*, and *Saccharomyces* extracts as a function of observed concentration (mM in the NMR tube; $N = 94$). Bars indicate the total number of metabolite occurrences in the three extracts within consecutive 2 mM windows, and the regression line indicates the best-fit power regression ($y = 42x^{-1.36}$, $r^2 = 0.86$). The internal standard (HEPES; $N = 3$), overlapped metabolites (acetate, lactate, and putrescine; $N = 9$), and nine metabolites with concentrations over 30 mM were excluded from the analysis. Metabolites observed in high-resolution $^1H-^{13}C$ HSQC spectra that were too dilute for FMQ by NMR ($N = 13$) were included with the quantifiable metabolites in the 0−2 mM bar. A complete list of all observed concentrations is provided in Supporting Information Table 1.

ing concentrations from three standard concentrations (2, 5, and 10 mM) is sufficient for estimating concentrations between 1 and 30 mM.

We recognize that quantitative estimates are influenced by any variable that affects spectrometer sensitivity (primarily salt concentration and NMR line shape). HEPES (5 mM) was included in all the samples to act as an internal control for these variables. HEPES is a convenient choice for 2D $^1H-^{13}C$ NMR studies, because it is a pH indicator and has several well-dispersed peaks (Figure 1). An additional benefit of HEPES normalization is that it allows standard curves generated for one study to be used in others, provided that all of the NMR instrument settings are identical. However, we found that HEPES normalization in the synthetic samples led to greater (10.2%) technical error (Supporting Information Figure 2). In our experience, accurate quantification (<3% error) requires measurements of absolute peak intensities referenced to concentration standards collected at the same time as the test samples.

Sample Size Requirements for Fast Quantification. Fast 2D $^1H-^{13}C$ NMR quantification requires a substantial amount of starting material. This study required 50 mg of metabolite extract (from ~400 mg of tissue) to generate signals that could be quantified in 12 minutes. The sample requirement can be reduced to 25 mg by using 5 mm susceptibility matched NMR tubes (e.g., from Shigemi, Inc.), which require smaller volumes. Smaller amounts of starting material will result in loss of information about metabolites present below the detection threshold. Longer NMR data collection times can partially compensate for lower concentra-

tions, but this approach is limited, because NMR sensitivity scales with the square root of the number of scans.

DISCUSSION

Our data indicate that 2D $^1H-^{13}C$ HSQC is superior to 1D ^1H NMR for quantitative analyses of complex mixtures. Although lower error estimates (~1%) have been reported for 1D ^1H NMR analyses,[17−19] these studies have been limited to well-dispersed peaks. One-dimensional ^1H NMR spectra of unfractionated biological extracts contain hundreds of overlapping resonances. Our analysis indicates that peak overlap considerably increases technical error.

We believe that the error estimates reported here represent the practical limit of quantitative precision for 1D ^1H NMR-based metabolomics. Our synthetic mixtures included only 26 metabolites, and we had the luxury of knowing the exact chemical shifts of every molecule from 2D $^1H-^{13}C$ HSQC spectra. NMR-based analyses of real biological samples must contend with >50 observable compounds with ambiguous resonance assignments and larger variations in chemical shift than our synthetic mixtures. Although these present serious obstacles for 1D ^1H NMR analysis, they can be overcome using the 2D $^1H-^{13}C$ NMR protocol we introduce here.

FMQ by NMR requires that target metabolites be identified prior to quantitative analysis. Although the identification of metabolites is laborious, it enables well-controlled quantitative analyses to be based on intensities of minimally overlapping peaks. Our experimental method and guidelines for reducing NMR acquisition times should make quantitative metabolite analyses feasible for biological studies in which 50-mg (extract dry mass) samples can be obtained.

SUPPORTING INFORMATION AVAILABLE

Table with the concentration of targets in the 24 synthetic mixtures analyzed. Tabulation of NMR peaks detected and assigned in the three cell extracts studied. Table with the experimentally determined concentrations of metabolites in the three biological systems. Illustration of a quantitative $^1H-^{13}C$ HSQC spectrum. Figure showing the statistical analyses carried out. This material is available free of charge via the Internet at http://pubs.acs.org.

(17) Malz, F.; Jancke, H. *J. Pharm. Biomed. Anal.* **2005**, *38*, 813−823.
(18) Burton, I. W.; Quilliam, M. A.; Walter, J. A. *Anal. Chem.* **2005**, *77*, 3123−3131.
(19) Pinciroli, V.; Biancardi, R.; Visentin, G.; Rizzo, V. *Org. Process Res. Dev.* **2004**, *8*, 381−384.