

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/41396648>

Profiling Polar and Semipolar Plant Metabolites throughout Extraction Processes Using a Combined Solution-State and High-Resolution Magic Angle Spinning NMR Approach

ARTICLE *in* ANALYTICAL CHEMISTRY · MARCH 2010

Impact Factor: 5.64 · DOI: 10.1021/ac9019076 · Source: PubMed

CITATIONS

44

READS

48

3 AUTHORS:



[Yasuyo Sekiyama](#)

National Food Research Institute

22 PUBLICATIONS 370 CITATIONS

[SEE PROFILE](#)



[Eisuke Chikayama](#)

Niigata University of International and Info...

26 PUBLICATIONS 484 CITATIONS

[SEE PROFILE](#)



[Jun Kikuchi](#)

RIKEN

154 PUBLICATIONS 3,144 CITATIONS

[SEE PROFILE](#)

Profiling Polar and Semipolar Plant Metabolites throughout Extraction Processes Using a Combined Solution-State and High-Resolution Magic Angle Spinning NMR Approach

Yasuyo Sekiyama,[†] Eisuke Chikayama,[†] and Jun Kikuchi^{*,†,‡,§}

RIKEN Plant Science Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 235-0045, Graduate School of Nanobiosciences, Yokohama City University, 1-7-29 Suehirocho, Tsurumi-ku, Yokohama 230-0045, and Graduate School of Bioagricultural Sciences and School of Agricultural Sciences, Nagoya University, 1 Furo-cho, Chikusa-ku, Nagoya-shi 464-8601, Japan

In metabolomic analyses, care should be exercised as to which metabolites are extracted from the sample and which remain in the residue; the remaining metabolites are typically discarded following the extraction process. In this study, nuclear magnetic resonance (NMR)-based metabolomics was used to visualize plant metabolite profiles throughout a series of repeated extraction processes. Metabolites remaining in the extraction residues of ¹³C-labeled *Arabidopsis thaliana* were recovered by repeated extraction using methanol-d₄ and deuterium oxide. The soluble extracts and residual pellets from each step of the extraction process were analyzed by both solution-state and high-resolution magic angle spinning NMR. Metabolic profiling based on chemical shifts in two-dimensional ¹H–¹³C heteronuclear single-quantum coherence spectra allowed the elucidation of both structural and chemical properties. In addition to the metabolite profile, there appears to be a relationship between metabolite structure and behavior throughout the repeated extraction process. These approaches suggest that metabolites are not always extracted in a single step and that the distribution of metabolites in an extraction scenario cannot be predicted solely on the basis of solubility or polarity. The composition of all metabolites in cells influences the solubility of each metabolite; thus, particular attention should be paid because changes in only a portion of the metabolites could influence the entire metabolite profile in a solvent extract.

Metabolomics is an emerging field that focuses on the systematic, comprehensive, and simultaneous profiling of small-molecule metabolites in cells. An accurate snapshot of the metabolome requires the reliable extraction of metabolites from

biological matrixes.^{1–6} For pure organic compounds, the total solubility parameter and the constituent partial solubility parameter are widely used for the selection of appropriate solvents.^{7,8} However, the simultaneous measurement of *all* components in the metabolome using a single, high-throughput extraction step is often complicated by the range of chemical and physical properties of the metabolites and the enormous range of concentrations at which they exist.¹ Nuclear magnetic resonance (NMR) can provide structure-based information on a global pool of metabolites, not only in isotropic solvent extracts, but also in heterogeneous samples, such as intact tissues or extraction residues containing insoluble materials.^{9–12} By combining these advantages with uniform ¹³C-labeling, we have been developing a methodology for “multidimensional” NMR-based metabolomics.^{13–19}

Generally, plant metabolomes are considered to be much larger than those of other organisms.²⁰ Some of the reported extraction solvents that have been employed in NMR-based plant metabolomics are summarized in Supporting Information Table

- (1) Colquhoun, I. J. *J. Pestic. Sci.* 2007, 32, 200–212.
- (2) Beckonert, O.; Keun, H. C.; Ebbels, T. M. D.; Bundy, J. G.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. *Nat. Protoc.* 2007, 2, 2692–2703.
- (3) Fiehn, O.; Wohlgemuth, G.; Scholz, M.; Kind, T.; Lee, D. Y.; Lu, Y.; Moon, S.; Nikolau, B. *Plant J.* 2008, 53, 691–704.
- (4) Kruger, N. J.; Troncoso-Ponce, M. A.; Ratcliffe, R. G. *Nat. Protoc.* 2008, 3, 1001–1012.
- (5) Viant, M. R. *Mol. Biosyst.* 2008, 4, 980–986.
- (6) Viant, M. R. *Methods Mol. Biol.* 2008, 410, 137–150.
- (7) Govers, H. A. J. *J. Chem. Soc., Faraday Trans.* 1993, 89, 3751–3759.
- (8) Stefanis, E.; Panayiotou, C. *Int. J. Thermophys.* 2008, 29, 568–585.
- (9) Lindon, J. C.; Holmes, E.; Nicholson, J. K. *FEBS J.* 2007, 274, 1140–1151.
- (10) Fan, T. W. M.; Lane, A. N. *Prog. Nucl. Magn. Reson. Spectrosc.* 2008, 52, 69–117.
- (11) Lindon, J. C.; Nicholson, J. K. *Annu. Rev. Anal. Chem.* 2008, 1, 45–69.
- (12) Fan, T. W. M.; Bird, J. A.; Brodie, E. L.; Lane, A. N. *Metabolomics* 2009, 5, 108–122.
- (13) Kikuchi, J.; Shinozaki, K.; Hirayama, T. *Plant Cell Physiol.* 2004, 45, 1099–1104.
- (14) Kikuchi, J.; Hirayama, T. *Methods Mol. Biol.* 2007, 358, 273–286.
- (15) Sekiyama, Y.; Kikuchi, J. *Phytochemistry* 2007, 68, 2320–2329.
- (16) Akiyama, K.; Chikayama, E.; Yuasa, H.; Shimada, Y.; Tohge, T.; Shinozaki, K.; Hirai, M. Y.; Sakurai, T.; Kikuchi, J.; Saito, K. *In Silico Biol.* 2008, 8, 339–345.
- (17) Chikayama, E.; Suto, M.; Nishihara, T.; Shinozaki, K.; Kikuchi, J. *PLoS ONE* 2008, 3, e3805.
- (18) Fukuda, S.; Nakanishi, Y.; Chikayama, E.; Ohno, H.; Hino, T.; Kikuchi, J. *PLoS ONE* 2009, 4, e4893.

* To whom correspondence should be addressed. E-mail: kikuchi@psc.riken.jp.

[†] RIKEN Plant Science Center.

[‡] Yokohama City University.

[§] Nagoya University.

S1^{3,4,19,21–36} (see also supplementary references S1–S15, Supporting Information). Although most of these extraction protocols have been elaborated to simultaneously quench metabolism and quantitatively extract the metabolites, attention should be focused on which metabolites are extracted from the sample and which remain in the residue. The main purpose of this study was to evaluate the multiple metabolites that can be extracted or left behind in the residue to improve our knowledge of plant metabolite extraction. Metabolites remaining in the extraction residues of ¹³C-labeled *Arabidopsis thaliana* were recovered by repeated extraction with methanol-*d*₄ (MeOD) and deuterium oxide (D₂O), two of the most common solvents in NMR-based metabolomics (see Table S-1). Changes in the metabolite profile in both the liquid extract and extraction residue were monitored after each extraction step using solution-state and high-resolution magic angle spinning (HR-MAS)³⁷ NMR. Statistical and structural metabolic profiling was performed throughout the extraction process, based on two-dimensional ¹H–¹³C heteronuclear single-quantum coherence (HSQC) spectra,^{38,39} which allowed for the determination of the relationship between the behavior and structural properties of the metabolites. These results will aid in our knowledge of *Arabidopsis* polar and semipolar metabolite extraction.

EXPERIMENTAL SECTION

Chemicals. MeOD (99.8% D) and HEPES-*d*₁₈ (98% D) were purchased from Cambridge Isotope Laboratories (Andover,

- (19) Mochida, K.; Furuta, T.; Ebana, K.; Shinozaki, K.; Kikuchi, J. *BMC Genomics* **2009**, *10*, e568.
- (20) Sumner, L. W.; Mendes, P.; Dixon, R. A. *Phytochemistry* **2003**, *62*, 817–836.
- (21) Ward, J. L.; Harris, C.; Lewis, J.; Beale, M. H. *Phytochemistry* **2003**, *62*, 949–957.
- (22) Moing, A.; Maucourt, M.; Renaud, C.; Gaudillere, M.; Brouquisse, R.; Lebouteiller, B.; Gousset-Dupont, A.; Vidal, J.; Granot, D.; Denoyes-Rothan, B.; Lerceteau-Kohler, E.; Rolin, D. *Funct. Plant Biol.* **2004**, *31*, 889–902.
- (23) Fan, T. W. M.; Higashi, R. M.; Lane, A. N.; Jardetzky, O. *Biochim. Biophys. Acta* **1986**, *882*, 154–167.
- (24) Cho, S. K.; Yang, S. O.; Kim, S. H.; Kim, H.; Ko, J. S.; Riu, K. Z.; Lee, H. Y.; Choi, H. K. *J. Pharm. Biomed. Anal.* **2009**, *49*, 567–571.
- (25) Manetti, C.; Bianchetti, C.; Casciani, L.; Castro, C.; Di Cocco, M. E.; Miccheli, A.; Motto, M.; Conti, F. J. *Exp. Bot.* **2006**, *57*, 2613–2625.
- (26) Le Gall, G.; Colquhoun, I. J.; Defernez, M. *J. Agric. Food Chem.* **2004**, *52*, 692–700.
- (27) Tarachiwin, L.; Ute, K.; Kobayashi, A.; Fukusakii, E. *J. Agric. Food Chem.* **2007**, *55*, 9330–9336.
- (28) Fortes, A. M.; Santos, F.; Choi, Y. H.; Silva, M. S.; Figueiredo, A.; Sousa, L.; Pessoa, F.; Santos, B. A.; Sebastiana, M.; Palme, K.; Malho, R.; Verpoorte, R.; Pais, M. S. *BMC Genomics* **2008**, *9*, e445.
- (29) Le Gall, G.; Puaud, M.; Colquhoun, I. J. *J. Agric. Food Chem.* **2001**, *49*, 580–588.
- (30) Defernez, M.; Colquhoun, I. J. *Phytochemistry* **2003**, *62*, 1009–1017.
- (31) Fumagalli, E.; Baldoni, E.; Abbruscato, P.; Piffanelli, P.; Genga, A.; Lamanna, R.; Consonni, R. *J. Agron. Crop Sci.* **2009**, *195*, 77–88.
- (32) Tarpley, L.; Duran, A. L.; Kebrom, T. H.; Sumner, L. W. *BMC Plant Biol.* **2005**, *5*, e8.
- (33) Choi, H. K.; Choi, Y. H.; Verberne, M.; Lefebvre, A. W. M.; Erkelens, C.; Verpoorte, R. *Phytochemistry* **2004**, *65*, 857–864.
- (34) Choi, Y. H.; Kim, H. K.; Linthorst, H. J. M.; Hollander, J. G.; Lefebvre, A. W. M.; Erkelens, C.; Nuzillard, J. M.; Verpoorte, R. *J. Nat. Prod.* **2006**, *69*, 742–748.
- (35) Tarachiwin, L.; Masako, O.; Fukusaki, E. *J. Agric. Food Chem.* **2008**, *56*, 5827–5835.
- (36) Toole, G. A.; Barron, C.; Le Gall, G.; Colquhoun, I. J.; Shawry, P. R.; Mills, E. N. C. *Planta* **2009**, *229*, 667–680.
- (37) Andrew, E. R.; Bradbury, A.; Eades, R. G. *Nature* **1959**, *183*, 1802–1803.
- (38) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185–189.
- (39) Lewis, I. A.; Schommer, S. C.; Hodis, B.; Robb, K. A.; Tonelli, M.; Westler, W. M.; Sussman, M. R.; Markley, J. L. *Anal. Chem.* **2007**, *79*, 9385–9390.

MA), deuterium oxide (99% D) was obtained from Sigma-Aldrich (St. Louis, MO), and a gas mixture containing 31% O₂, 79% N₂, and the desired ¹³CO₂ concentration (345–350 ppm by volume) was purchased from Takachiho Chemical Industrial (Tokyo, Japan).

Preparation of NMR Samples. Uniform Stable Isotope Labeling of *A. thaliana*. *A. thaliana* (ecotype Columbia 0) was germinated and grown on rock wool blocks (25 seeds per block; Nittobouseki, Tokyo, Japan) at 22 °C under a 14 h light (8:00 a.m. to 10:00 p.m.)/10 h dark (10:00 p.m. to 8:00 a.m.) cycle at a light intensity of 90 μmol/(m² s) in a growth chamber (NK System Biotron; Nippon Medical and Chemical Instrument, Osaka, Japan). The plants were thinned out to three or four per block after 3 weeks. A gas mixture containing 31% O₂, 79% N₂, and 345–350 ppm by volume ¹³CO₂ was passed continuously through the chamber from the germination stage until the early flowering stage (35 days). Nutrient solution (1000-fold-diluted HYPONeX 6:10:5; HYPONeX Japan, Osaka, Japan) was supplied to each block of rock wool once a week until the end of the experiment. The aerial components of about 30 plants were collected at 2:00 p.m. on day 35.

Extraction. The aerial components of labeled *A. thaliana* were frozen immediately in liquid nitrogen and lyophilized in a VO-800F freeze-dryer (Taitec, Cupertino, CA). All dried materials were combined and ground in an Auto-Mill TK AM4 (Tokken, Chiba, Japan) to yield 690 mg of stock powder, which was stored at –80 °C until extraction. All extraction experiments in this study were performed using 20 mg aliquots of the stock powder; i.e., the data represent the average of 30 plants obtained from the labeling experiment, performed once. For extraction, ¹³C-*A. thaliana* powder (20 mg) was suspended in 500 μL of MeOD or in 500 μL of D₂O. The mixture was heated at 50 °C for 5 min in a Thermomixer comfort (Eppendorf AG, Hamburg, Germany) and then centrifuged (10000g, 5 min). To maintain the concentration of sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS), MeOD (including 0.5 mM DSS) and D₂O (including 1.0 mM DSS) were used for extraction when the resulting residue was not used in HR-MAS analyses or further extractions. Solution-state NMR experiments were performed with the supernatant. The pellets, which contained insoluble materials, were lyophilized and analyzed by HR-MAS or extracted repeatedly with MeOD or D₂O according to the scheme in Figure 1. Residual solvent was removed by lyophilization prior to the next extraction step. The samples analyzed by NMR are listed in Table 1. The pH (or pD) values of the supernatant of three independent extractions were checked using a calibrated Horiba twinpH B-212 (Horiba, Kyoto, Japan). Calibration of the pH meter was performed using aqueous standard solutions, phosphate buffer (pH 6.86 at 25 °C), and phthalate buffer (pH 4.01 at 25 °C), purchased from Horiba. The supernatant pH (or pD) values were 5.8, 5.9, and 5.8 (M1), 6.0, 5.9, and 6.0 (M2), 5.9, 5.9, and 6.0 (M3), 5.8, 5.8, and 5.8 (W1), 6.2, 6.2, and 6.1 (M1W1), 6.5, 6.5, and 6.5 (M3W1), and 7.0, 7.0, and 7.0 (M3W2). Although accurate determinations of pH (or pD) in organic solvents are difficult, it was at least suggested that no extreme change in pH value was observed among M1, M2, and M3. HSQC analyses of three independent extractions were performed on the first MeOD extract (M1 in Figure 1), D₂O extract (W1 in Figure 1), and D₂O extract of

the M1 pellet (M1W1 in Figure 1); error estimates are shown in Tables S4–S6, Supporting Information. The ratio of the standard deviation to the average signal intensity (SD/average) varied from 0.004 (P1041, sucrose) to 0.221 (P2261, proline) in M1, from 0.000 (P234, glucose) to 0.403 (P84, uridine) in W1, and from 0.015 (P1014, glycerol) to 0.600 (P264, ascorbate) in M1W1. The degree of error increased with the number of extraction steps. The observed reproducibility of most of the peaks allows a discussion of metabolite behavior throughout the extraction process. The effects of the extraction temperature and time, provided in the accompanying Supporting Information, were examined by three independent D₂O (including 1.0 mM DSS, 500 μL) extractions of ¹³C-*A. thaliana* (5 mg of stock powder) at 50 and 90 °C for 1, 5, and 30 min. No remarkable differences were observed in the resulting ¹H NMR spectra (Supporting Information Figure S-1). The stability of the samples prepared by the procedure described above was checked by analyzing samples kept in the dark at 298 K for 7 h (in the NMR magnet for the ¹H–¹³C HSQC measurement) followed by –30 °C (sample storage) for several days over the period allowed for the experiment (Supporting Information Figure S-2).

NMR Analyses. Solution NMR. Sample solutions were transferred into 5 mm Ø NMR tubes. All NMR spectra were recorded on an Avance-700 spectrometer (Bruker, Billerica, MA) equipped with an inverse triple-resonance CryoProbe with a z-axis gradient for 5 mm sample diameters operating at 700.153 MHz for ¹H and 176.061 MHz for ¹³C. The temperature of all NMR samples was maintained at 298 K. The chemical shifts were referenced to the TMS group of DSS (0.5 mM in MeOD, 1.0 mM in D₂O) internal standard.

HR-MAS Analysis. An aliquot of the lyophilized pellet (7 mg) was soaked in 50 μL of K₂HPO₄/KH₂PO₄ buffer (100 mM in D₂O, including 1.0 mM DSS, pH 7.0) and placed in a zirconium oxide rotor with a 4 mm outer diameter. ¹H–¹³C HSQC spectra were acquired on an AVANCE-500 spectrometer (Bruker) equipped with a g-HR MAS 500 SB BL4 probe operating at 500.132 MHz for ¹H and 125.764 MHz for ¹³C. The samples were spun at 8 kHz and maintained at 298 K throughout the experiment. The chemical shifts were referenced to the TMS group of DSS (1.0 mM in KPi) internal standard.

Two-dimensional ¹H–¹³C HSQC Spectroscopy. Two-dimensional ¹H–¹³C HSQC spectra were collected using echo/antiecho gradient selection (hsqcetgp pulse program in the Bruker library). The proton 90° pulse value was checked and set to each sample (typically 10 μs for MeOD extracts, 11 μs for D₂O extracts, and 10 μs for HR-MAS samples). In all cases, the length of the carbon 90° pulse was 15.0 μs. Relaxation delay values of 3.0 and 2.0 s were used in solution-state and HR-MAS analysis, respectively. In total, 256 complex f1 (¹³C) and 1024 complex f2 (¹H) points were recorded with 32 scans (solution NMR analysis) or 64 scans (HR-MAS analysis) per f1 increment. The spectral window and offset frequency in the f1 dimension were 76 and 48 ppm, respectively. The spectral window in the f2 dimension was 16 ppm. The offset frequency in the f2 dimension was 3.3 ppm for MeOD extracts or 4.8 ppm for D₂O extracts.

Data Processing. For metabolic profiling, the spectra were processed with NMRPipe and analyzed using NMRDraw.⁴⁰ All

spectra were Fourier-transformed with the Lorentzian-to-Gaussian window function (10 Hz Lorentzian line width and 15 Hz Gaussian line width for the f1 dimension; 5 Hz Lorentzian line width and 10 Hz Gaussian line width for the f2 dimension). Data were zero-filled to 1024 (f1) and 4096 (f2) data points. An automatic polynomial baseline correction was subsequently applied in both dimensions, and a linear prediction was applied to the f1 dimension.

Chemical-Shift-Based Profiling of ¹H–¹³C HSQC Spectra. Peak intensities in the ¹H–¹³C HSQC spectra were standardized against the calculated noise level of each spectrum. Each ¹H–¹³C HSQC spectrum was divided into 76 × 50 grids (Supporting Information Figure S-5), and 460 grids containing peaks were used for the analysis. The sum of the standardized signal intensities within each grid was used to calculate the Z-scores used for hierarchical clustering as the (sum of the signal height in each grid – average of all grids in the same spectra)/(standard deviation of all grids in the same spectra).

The Z-scores were applied directly to a hierarchical clustering⁴¹ and heat map display using software developed at Stanford University.^{42,43} The cluster algorithm was set to complete linkage clustering using Pearson's correlation.

Collection of Standard Chemical Shifts. Standard chemical shifts for aqueous samples were collected using KPi buffer (pH 7.0) as described previously.¹⁷ Standard chemical shifts for MeOD extracts were collected using MeOD/HEPES-d₁₈ buffer. The MeOD/HEPES-d₁₈ buffer was prepared by mixing a 95:5 (v/v) ratio of MeOD and 200 mM HEPES-d₁₈ solution in D₂O (adjusted to pH 7.0 with NaOD in D₂O) and DSS (final concentration 0.5 mM), affording MeOD/HEPES (10 mM) solution. The final pH of the MeOD/HEPES solution was verified between 6.8 and 6.9 as described above for the preparation of NMR samples.

Annotation of Signals for Metabolic Profiling. Common peak ID numbers were given to corresponding HSQC signals throughout the MeOD and aqueous series (D₂O extracts and pellet in KPi), respectively. A common peak ID then represents the same signal throughout the same solvent series. With signal overlap, however, a common peak ID can represent one or more metabolites. Candidate metabolites for each common peak ID were determined by examining four factors: the degree of match between the chemical shift of the signals and that of standard compounds, the ratio of the number of signals in a given spectrum to the theoretical number of standard compounds, χ² analysis, and Pearson's correlation analysis. Candidate metabolites for the peaks observed in each spectrum were selected from standard compounds by comparing the chemical shift difference. A compound was selected when the chemical shift difference between the standard and queried peak was less than 0.03 and 0.53 ppm for ¹H and ¹³C,

-
- (40) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.
 - (41) Cheadle, C.; Vawter, M. P.; Freed, W. J.; Becker, K. G. *J. Mol. Diagn.* **2003**, *5*, 73–81.
 - (42) Eisen, M. B.; Spellman, P. T.; Brown, P. O.; Botstein, D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14863–14868.
 - (43) Saeed, A. I.; Sharov, V.; White, J.; Li, J.; Liang, W.; Bhagabati, N.; Braisted, J.; Klapa, M.; Currier, T.; Thiagarajan, M.; Sturn, A.; Snuffin, M.; Rezantsev, A.; Popov, D.; Ryltsov, A.; Kostukovich, E.; Borisovsky, I.; Liu, Z.; Vinsavich, A.; Trush, V.; Quackenbush, J. *Biotechniques* **2003**, *34*, 374–378.

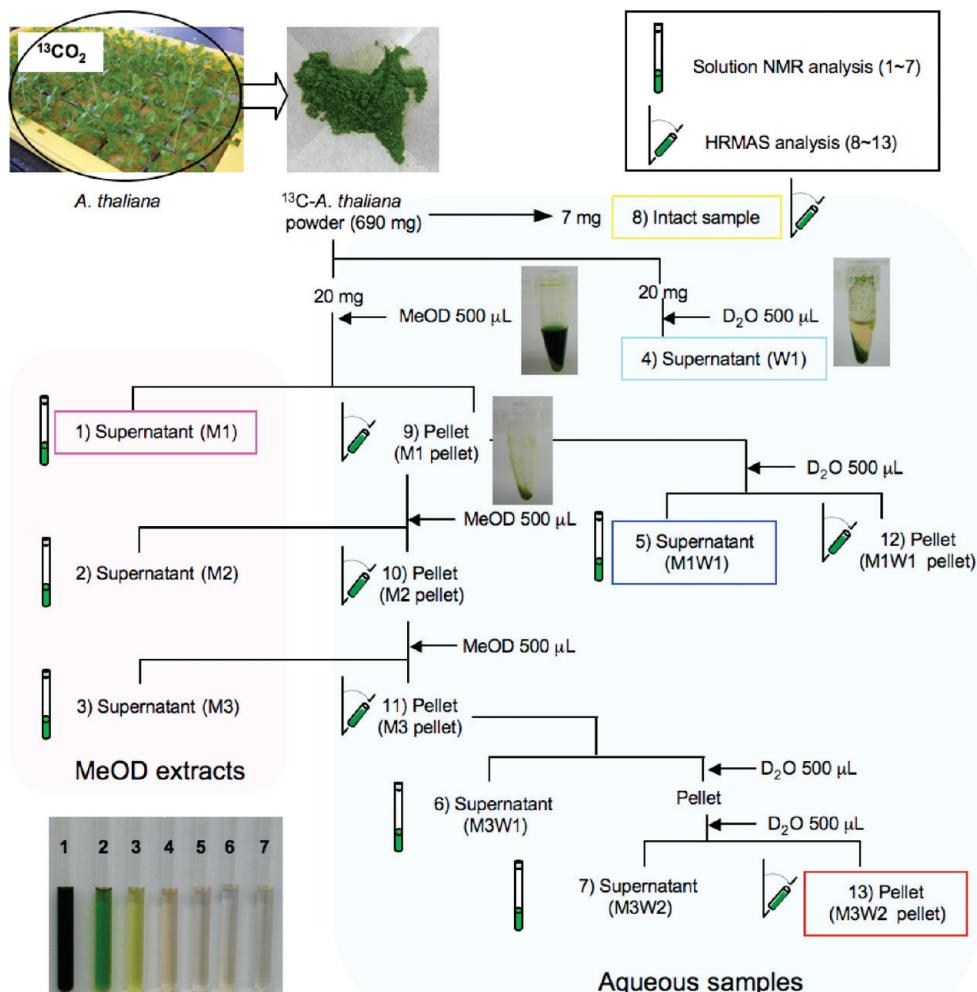


Figure 1. Sequential extraction scheme for ^{13}C -*A. thaliana*. The colors of the rectangles correspond to those of the NMR peaks in Figure 2. The samples of solvent extract are shown in the lower left.

respectively. If several compounds were selected for one signal, each compound was ranked in order of the ratio of the observed number of signals to the theoretical number of signals for that metabolite, and the compounds with the highest ratio were selected as temporary candidates for the queried peak.

Next, a χ^2 analysis was used to evaluate chemical shift fluctuations. The χ^2 value of each peak was defined as

$$\chi^2 = \sum D/2 \left[\left\{ \frac{(^1\text{H} \text{ chemical shift}_{\text{obsd}} - ^1\text{H} \text{ chemical shift}_{\text{std}})^2}{\sigma_{\Delta\text{H}}^2} \right\} + \left\{ \frac{(^{13}\text{C} \text{ chemical shift}_{\text{obsd}} - ^{13}\text{C} \text{ chemical shift}_{\text{std}})^2}{\sigma_{\Delta\text{C}}^2} \right\} \right]$$

where $\sigma_{\Delta\text{H}}$ and $\sigma_{\Delta\text{C}}$ are the standard deviations of the differences in the ^1H and ^{13}C chemical shifts, respectively, and D is the degrees of freedom for the corresponding χ^2 distribution, which equaled the number of spectra in which the queried peak was observed. χ^2 values were calculated independently for the MeOD and aqueous series. The candidate compounds were rearranged to minimize the area under the right-hand tail of the χ^2 distribution. The annotated metabolites and chemical

shift data are listed for the aqueous and MeOD series in Supporting Information Tables S-2 and S-3, respectively. Pearson's correlation analysis between peaks annotated to the same metabolite was performed to evaluate whether the intensities of these peaks were synchronized throughout repeated extraction processes (correlation coefficients are listed in Supporting Information Figures S-3 and S-4).

In a final step, the assignment of each carbon atom in a candidate metabolite was predicted by comparing the spectra with those of structurally related compounds in our database or in public NMR databases, such as the Biological Magnetic Resonance Data Bank (BMRB; <http://www.bmrb.wisc.edu>)⁴⁴ and the Human Metabolome Database (HMDB; <http://www.hmdb.ca>).⁴⁵ The carbon atoms in the candidate metabolites were numbered

- (44) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Wenger, R. K.; Yao, H. Y.; Markley, J. L. *Nucleic Acids Res.* **2008**, *36*, D402–D408.
- (45) Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S.; Mandal, R.; Sinelnikov, I.; Xia, J. G.; Jia, L.; Cruz, J. A.; Lim, E.; Sobsey, C. A.; Shrivastava, S.; Huang, P.; Liu, P.; Fang, L.; Peng, J.; Fradette, R.; Cheng, D.; Tzur, D.; Clements, M.; Lewis, A.; De Souza, A.; Zuniga, A.; Dawe, M.; Xiong, Y. P.; Clive, D.; Greiner, R.; Nazyrova, A.; Shaykhutdinov, R.; Li, L.; Vogel, H. J.; Forsythe, I. *Nucleic Acids Res.* **2009**, *37*, D603–D610.

Table 1. List of NMR Samples, Abbreviations, and Observed Numbers and Intensities of the Signals in ^1H – ^{13}C HSQC Spectra^a

no.	sample	abbrev	material for extraction	material mass (mg)	experiment	residue mass (mg)	number of signals	total peak intensity
1	first MeOD extract	M1	intact sample	20.0	solution NMR	12.0	343	880.3
2	second MeOD extract	M2	M1 pellet	12.0	solution NMR	12.0	210	112.1
3	third MeOD extract	M3	M2 pellet	12.0	solution NMR	12.0	79	41.4
4	first D ₂ O extract	W1	intact sample	20.0	solution NMR	12.6	295	304.7
5	D ₂ O extract of the pellet of the first MeOD extraction	M1W1	M1 pellet	12.0	solution NMR	9.1	220	181.1
6	D ₂ O extract of the pellet of the third MeOD extraction	M3W1	M3 pellt	12.0	solution NMR	8.7	212	133.1
7	D ₂ O extract of the pellet of the third MeOD extraction	M3W2	M3W1 pellet	12.0	solution NMR	9.7	44	14.3
8	intact sample	intact		7.0	HR-MAS		144	55.2
9	pellet of the first MeOD extraction	M1 pellet		6.5	HR-MAS		93	22.8
10	pellet of the second MeOD extraction	M2 pellet		6.5	HR-MAS		74	13.3
11	pellet of the third MeOD extraction	M3 pellet		7.0	HR-MAS		54	9.8
12	pellet of the first D ₂ O extraction of the pellet of the first MeOD extraction	M1W1 pellet		7.0	HR-MAS		20	1.8
13	pellet of the second D ₂ O extraction of the pellet of the third MeOD extraction	M3W2 pellet		7.0	HR-MAS		5	0.9

^a Peak intensities in the spectra were standardized against the calculated noise level of each spectrum.

successively on the basis of IUPAC guidelines. The annotations of organic acids were confirmed by mixing standard compounds (citric acid, malic acid, [$^{13}\text{C}_3$]sodium pyruvate, [$^{13}\text{C}_4$]succinic acid, [$^{13}\text{C}_4$]fumaric acid, [$^{13}\text{C}_2$]sodium acetate, and [$^{13}\text{C}_3$]sodium pyruvate) with the sample solutions.

Metabolic Profiling. For metabolic profiling, one peak was selected as representative of each metabolite on the basis of the χ^2 calculation and Pearson's correlation analysis. In this study, we postulated that the signals that showed high correlation coefficients of over 0.9 originated from the same metabolite. The signal that afforded a larger right-hand tail of the χ^2 distribution was selected from the signals of the same metabolite. The intensity of each annotated peak was standardized against the calculated noise level of each spectrum and normalized using the Z-score transformation across the spectra of 10 aqueous samples. The Z-scores were directly subjected to hierarchical clustering using the procedure described for chemical-shift-based profiling of ^1H – ^{13}C HSQC spectra.

RESULTS AND DISCUSSION

Profiling of ^1H – ^{13}C HSQC Spectra during Repeated Extraction. Samples of ^{13}C -labeled *A. thaliana*, from the same harvest, were extracted repeatedly using MeOD or D₂O until only a few signals could be detected in the two-dimensional ^1H – ^{13}C HSQC spectra, according to the scheme depicted in Figure 1. The labeling experiment was performed once, and aliquots of the resultant stock powder were used for extraction experiments. Briefly, ^{13}C -*A. thaliana* powder was extracted using MeOD or D₂O, and the residual pellets from each step were either analyzed directly by HR-MAS or extracted further to yield 13 samples (3 MeOD extracts and 4 D₂O extracts for solution-state analysis, 1 intact sample, and 5 residual pellets for HR-MAS analysis). The samples and abbreviations used for the NMR analyses and the observed number and total intensity of

peaks in the ^1H – ^{13}C HSQC spectra are listed in Table 1. HSQC analyses of three independent extractions were performed on the first MeOD extract (M1 in Figure 1), D₂O extract (W1 in Figure 1), and D₂O extract of the M1 pellet (M1W1 in Figure 1), and the technological error was estimated (Supporting Information Tables S4–S6).

An overlay of five representative ^1H – ^{13}C HSQC spectra (M1, W1, M1W1, intact, and M3W2 pellet) with representative peak ID numbers is shown in Figure 2. An aliphatic region (δ_{H} 0.6–1.5 ppm), including fatty acids, was abundant in all three MeOD extracts (M1–3). The D₂O extracts (the direct D₂O extract of the intact sample (W1) and subsequent D₂O extracts of the MeOD extraction residues (M1W1–M3W2)) were rich in hydrophilic metabolites (δ_{H} 3.4–4.0 ppm), such as carbohydrates. The HR-MAS spectra of the intact plant powder and the residual pellets after each extraction step (M1 pellet–M3W2 pellet), which were measured in aqueous buffer, were similar to those of the D₂O extracts. Detailed profiles of the observed HSQC signals are shown in Figure 3 using hierarchical clustering based on the chemical shifts and peak intensities (see also Supporting Information Figure S5). Samples from the repeated extractions can be classified roughly as MeOD extracts or aqueous samples (D₂O extracts and samples for HR-MAS analyses), depending on the relative abundance of aliphatic and hydrophilic regions, as described above. The detectable metabolites are likely biased to some extent, depending on the polarity of the extraction solvent.^{7,8} Thus, a combination of polar and nonpolar solvents was used to afford a more comprehensive extraction of metabolites (Supporting Information Table S-1).

Significant amounts of metabolites remained in the residues after a single extraction step (Table 1). Several signals were still observed in the M3W2 pellet, which represents the residue after five repeated extraction steps with MeOD and D₂O. Upon careful

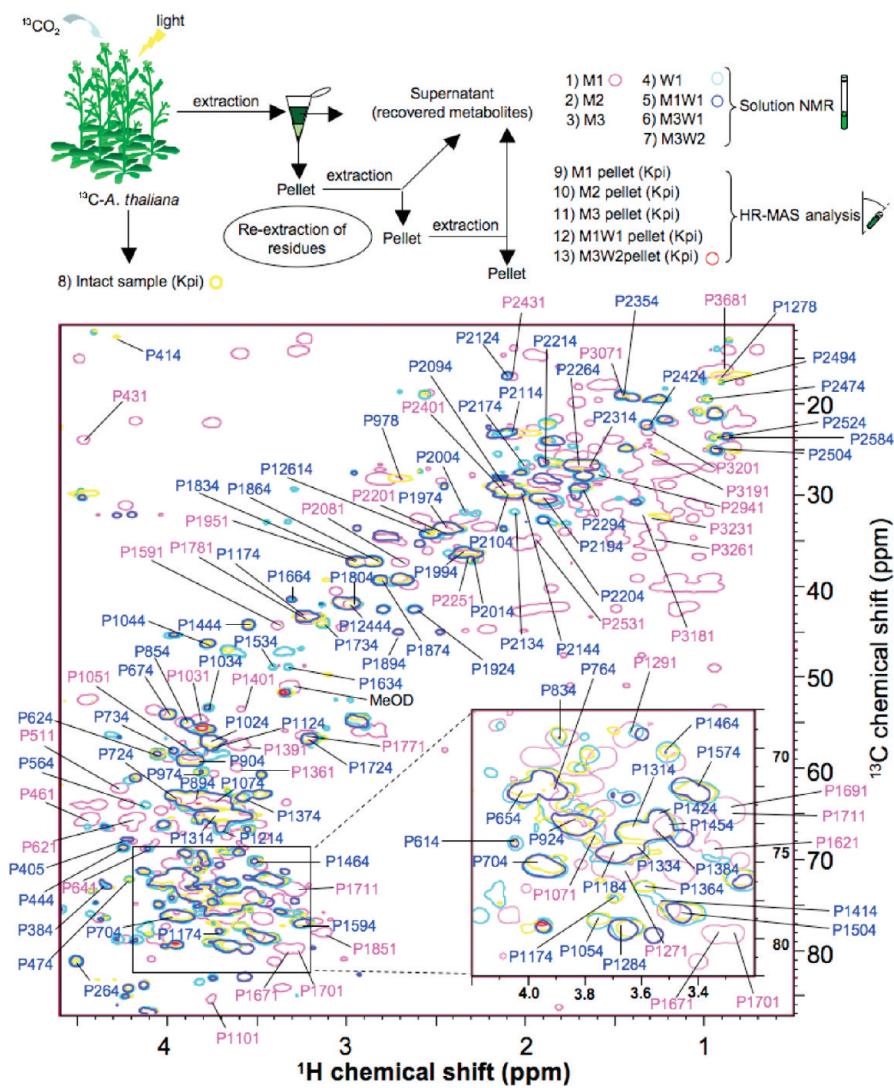


Figure 2. Overlaid ^1H - ^{13}C HSQC spectra of ^{13}C -A. thaliana extracts (analyzed by solution NMR) and residual pellets (analyzed by HR-MAS). The spectral window and offset frequency in the f1 dimension were 76 and 48 ppm, respectively. The spectral window in the f2 dimension was 16 ppm. The offset frequency in the f2 dimension was 3.3 ppm for MeOD extracts and 4.8 ppm for D₂O extracts and residual pellets. Peak ID numbers in pink indicate MeOD extracts; those in blue denote aqueous samples (D₂O extracts and residual pellets).

examination of the HSQC spectra, shown in Figure 3, it was revealed that the relative peak intensities among the grids within a spectrum were not completely maintained during the extraction processes in which the same solvent was applied repeatedly. This result indicates that the behavior of a metabolite under an extraction scenario is context-dependent. Thus, detailed metabolic profiling was performed to examine the behavior of each metabolite throughout the extraction process.

Profiling Metabolites in Extraction Residues. To shed light on the behavior of metabolites during the extraction process, we focused on the components that were not extracted in a single step, but remained in the extraction residue. Information on these metabolites was obtained from the 10 aqueous samples (i.e., the D₂O extracts (W1–M3W2), which included metabolites recovered from MeOD extraction residues, the intact sample, and the residual pellets (M1 pellet–M3W2 pellet)). To obtain structural information, candidate metabolites for each peak in the HSQC spectra were identified by matching the signals in the spectra with chemical shifts of standard compounds, performing χ^2 calculations, and applying Pearson's correlation

analysis. The annotated peaks and their intensities are listed in Supporting Information Tables S-2 and S-3. Although the standard chemical shifts were collected at pH 7.0 in KPi for aqueous samples and pH 6.9 in MeOD/HEPES solution for MeOD extracts, the pH values of the NMR samples were confirmed to be 5.8–6.5. However, pH-dependent chemical shift fluctuations of most metabolites should be expected to be less than $\Delta\delta_H$ 0.03 ppm and $\Delta\delta_C$ 0.53 ppm (tolerances used for selecting candidate metabolites; see the Experimental Section) within the range of the observed pH.¹⁷ Organic acid annotations were confirmed by mixing standard compounds with sample solutions. A χ^2 analysis of the chemical shifts of peaks present in different spectra reduced annotation complexities resulting from chemical shift perturbations due to, for example, signal overlap and pH fluctuations. The whole of the ^{13}C pattern, which is unique for a given candidate molecule, can be considered in predictions of metabolite identity. Pearson's correlation analysis was used to compare the intensities of all detectable signals in different samples that were annotated as the same metabolite, and the reliability of the annotations was verified by examining whether

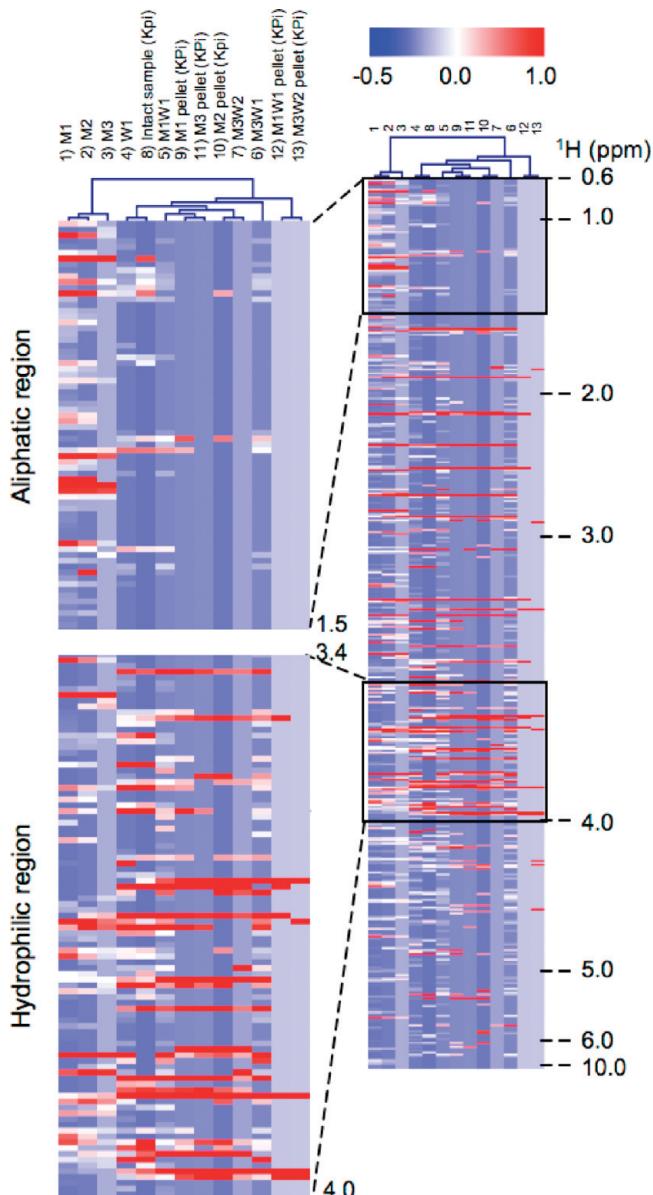


Figure 3. Clustered heat map showing the signal variation in the ^1H - ^{13}C spectra of ^{13}C -*A. thaliana* during repeated extraction processes. An expanded view of the aliphatic and hydrophilic regions is shown on the left. The columns represent different extracts or residual pellets, and the row represents each grid.

the intensities of these peaks were synchronized throughout the extraction process. Correlation coefficients between those peaks attributed to the same molecule are shown in Figures S-3 and S-4, Supporting Information. Low degrees of correlation may indicate peak overlap or that the peaks do not correspond to the same molecule.

Using this method, 41 metabolites were annotated in the aqueous samples. For metabolic profiling, one representative peak was selected for each metabolite on the basis of the χ^2 calculations and Pearson's correlation analysis. Although these annotations were not entirely unambiguous, the characteristic chemical shifts suggested that the annotated signals originated from carbons in similar partial structures or with neighboring functional groups common to the candidate metabolite. This advantage of NMR analysis allowed for the determination of metabolite profiles with additional structural and chemical

information. A clustered heat map of the peak intensities is illustrated in Figure 4A, and the raw intensity of each peak is plotted in Figure 4B. Hierarchical clustering was performed using the procedure described above for chemical-shift-based profiling of ^1H - ^{13}C HSQC spectra, except that the Z-score transformation of the intensity for each annotated peak was performed across spectra obtained from the 10 aqueous samples.

The annotated metabolites could be classified roughly into several clusters on the basis of their behavior during repeated extractions. The metabolites in cluster 1 were extracted consistently in repeated MeOD extractions, whereas cluster 2 metabolites remained in the residual pellets. Cluster 2 metabolites were further categorized into two groups: those extracted in modest amounts by MeOD (group 1) and those nearly insoluble in MeOD (group 2). The D₂O extraction efficiency of metabolites in group 2 was increased by the preceding MeOD extractions. Structures of the candidate metabolites are shown in Figure 5. The carbon atoms assigned to the queried peaks, which were used for profiling, are indicated with blue solid circles. Blue open circles indicate peaks with a high degree of correlation to the peaks used for profiling. Although it is difficult to show a high degree of correlation between all of the signals for a given molecule due to overlap with other metabolite(s), unambiguous annotation, or undetectable metabolite levels, the signals that exhibited a high degree of correlation should derive from the same molecule and can be used to predict structural properties. The colored open rectangles in Figure 5 highlight the information obtained from HR-MAS analyses. Although the relatively low sensitivity of HR-MAS resulted in fewer detectable metabolites and lower peak intensities, the relative intensity of each peak was retained throughout the extraction process. The intensities reflect the true yield of the metabolites from each sample with no bias caused by solvent extraction, enabling the estimation of the relative solubility of the metabolites in each cluster. The relative solubility of 17 metabolites, which could be annotated in both MeOD and D₂O extracts, are compared in Supporting Information Figure S-6. Most of these metabolites were extracted by both MeOD and D₂O, although no signal for glutamate (P1994) or aspartate (P854) was observed in the MeOD extracts. Figure 5 shows several relationships between the structures and yields of metabolites. Cluster 1 metabolites tended to include monosaccharides and nonpolar amino acids, which were extracted with MeOD. Conversely, cluster 2 metabolites included polar amino acids and other metabolites functionalized with multiple polar groups, such as phosphates, carboxyl groups, and amines, which were soluble in D₂O. The profiles of group 2 of cluster 2 suggest that the aqueous extraction yields of certain sugar phosphates, oligosaccharides, organic acids, and acidic amino acids were increased by the preceding removal of competitive metabolites that were soluble in both water and methanol. These results suggest that the yields and profiles of metabolites cannot be predicted by the polarity of the metabolites and solvents alone, as can be done with pure organic compounds. In these experiments, organic acids of central metabolism were not detected successfully. It is difficult to explain exactly why major organic acids such as citrate and fumarate, which have been reported to be detectable in *Arabidopsis* aqueous methanol or water extracts,^{21,39,52} were not detected in this study. In NMR

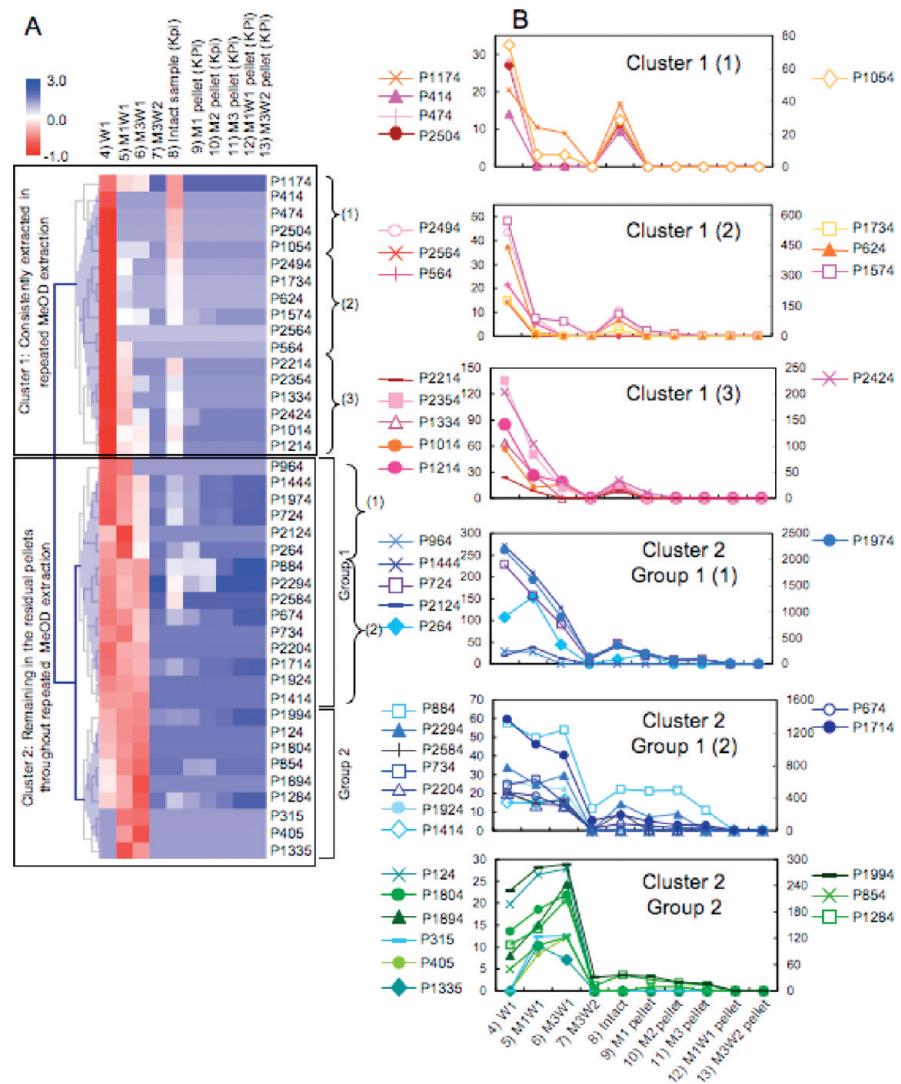


Figure 4. Profiles of the metabolite yields in D_2O extracts and residual pellets are shown using a clustered heat map (A) with the intensities of the annotated signals throughout the extraction process (B). In the heat map, the columns represent different extracts or insoluble pellets and the rows represent individual annotated peaks. The row peak intensities were normalized by Z-score transformation across the samples before hierarchical clustering.

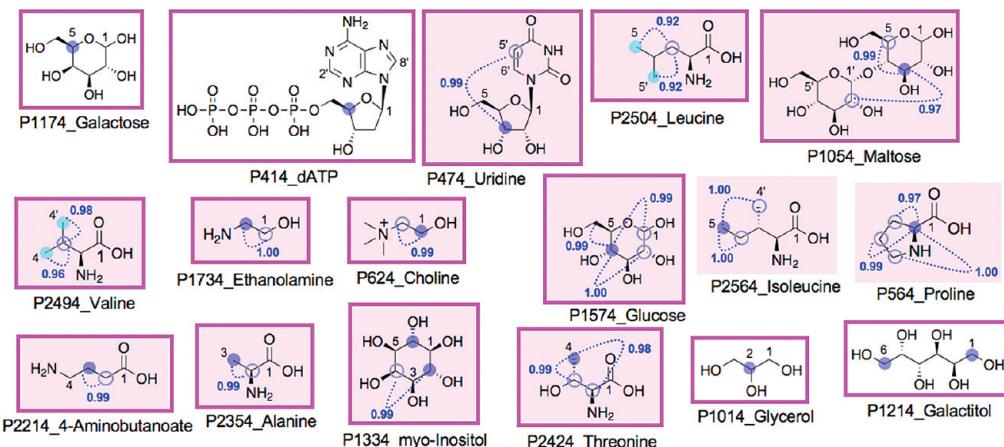
analysis, it is known that paramagnetic ions, such as Fe^{3+} , Cu^{2+} , and Mn^{2+} , cause problems in line broadening and may result in decreased signals or the disappearance of carboxylic acids. The removal of such paramagnetic ions by chelating agents or resins may improve the spectra.^{10,46,47} Nevertheless, the results described above would be useful in the consideration of *Arabidopsis* polar and semipolar metabolite extraction and interpretation of the obtained metabolite profiles.

- (46) Lane, A. N.; Fan, T. W. M.; Higashi, R. M. *Biophysical Tools for Biologists: Vol. 1 In Vitro Techniques*; Methods in Cell Biology, Vol. 84; Academic Press: San Diego, CA, 2008; p 541.
- (47) Lundberg, P.; Lundquist, P. O. *Planta* **2004**, *219*, 661–672.
- (48) Nakayama, F. S. *Ind. Crops Prod.* **2005**, *22*, 3–13.
- (49) Scordino, M.; Di Mauro, A.; Passerini, A.; Maccarone, E. *J. Agric. Food Chem.* **2005**, *53*, 651–658.
- (50) Alcaide, E. M.; Nefzaoui, A. *Int. Biodeterior. Biodegrad.* **1996**, *38*, 227–235.
- (51) Fan, W. M. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996**, *28*, 161–219.
- (52) Topcu, G.; Ulubelen, A. *J. Mol. Struct.* **2007**, *834*, 57–73.
- (53) Xi, Y. X.; de Ropp, J. S.; Viant, M. R.; Woodruff, D. L.; Yu, P. *Anal. Chim. Acta* **2008**, *614*, 127–133.
- (54) Weljie, A. M.; Newton, J.; Jirik, F. R.; Vogel, H. *J. Anal. Chem.* **2008**, *80*, 8956–8965.

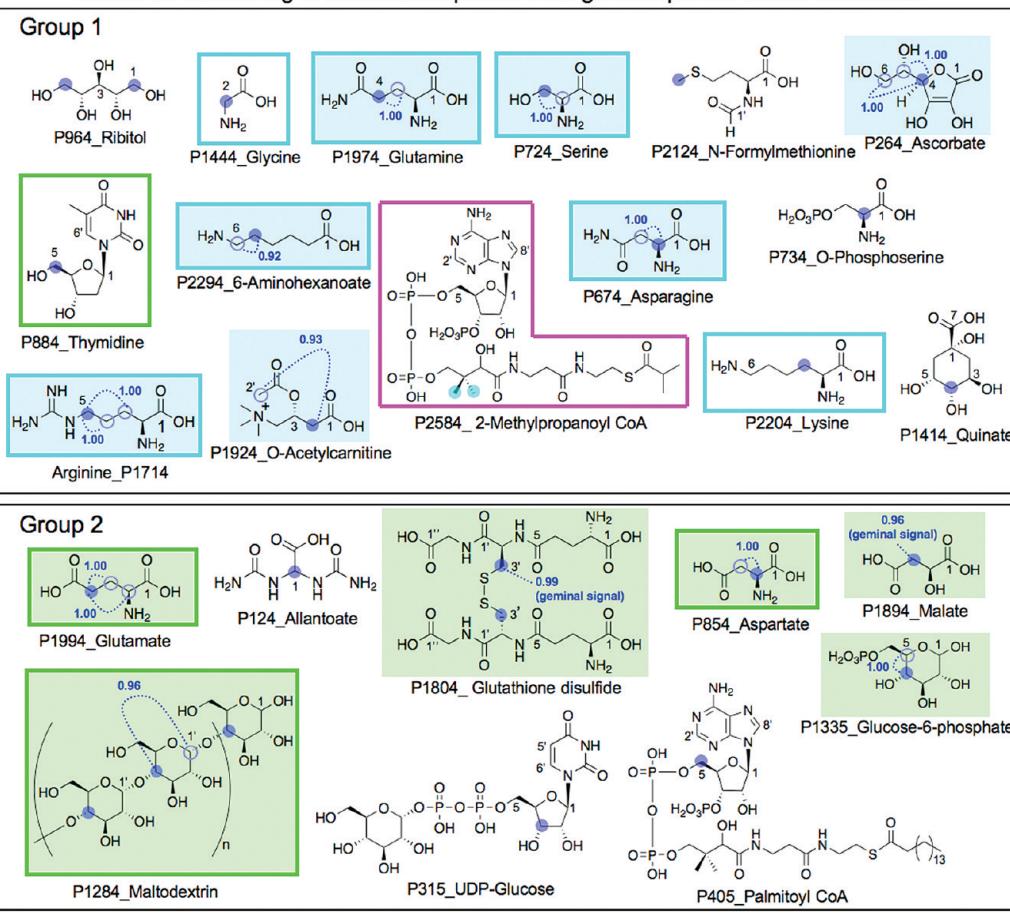
Advantage of Metabolic Profiling for the Evaluation of Extraction Processes. The results described above indicate that metabolite profiles were affected not only by the solubility of each metabolite, but also by the chemical properties and proportions of *all* the metabolites in the sample. Additionally, information on residual metabolites is important not only in conducting metabolic profiling, but also for future industrial applications of plant chemical resources.^{48–50} Plants are the source of a vast array of metabolites, such as amino acids, nucleic acids, carbohydrates, phenols, organic acids, fatty acids, and vitamins, and appropriate methods for evaluating both the extraction process and the obtained metabolite mixtures are required for supplying and recycling these chemical resources.

NMR spectroscopy offers a means of noninvasive and rigorous structural analysis of metabolites in crude extracts or intact tissues. The sensitivity and resolution of two-dimensional 1H – ^{13}C HSQC analyses enable improved peak identification and structural analysis of even unknown metabolites.^{10,51–54} Because many metabolites can be characterized on the basis of characteristic

Cluster 1: Consistently extracted in repeated MeOD extraction



Cluster 2: Remaining in the residual pellets throughout repeated MeOD extraction



- - Annotated signal used for profiling
- - Prochiral carbons for which the assignments are interchangeable
- - Signal which shows high correlation (correlation coefficient >0.9) to the signal used for profiling (●)
- - Extracted by MeOD (HRMAS analysis)
- - Modestly extracted by MeOD (HRMAS analysis)
- - Nearly insoluble in MeOD (HRMAS analysis)

Figure 5. Structures of candidate metabolites. The carbons assigned to queried peaks are indicated with blue solid circles. Light blue solid circles correspond to prochiral carbons for which the assignments are interchangeable. Blue open circles mark peaks with a high degree of correlation to the peaks used for profiling. Pearson's correlation coefficients between those peaks are shown in bold blue. Metabolites with high correlation coefficients are highlighted with pale pink (in cluster 1), pale blue (in group 1 of cluster 2), and pale green (in group 2 of cluster 2). The pink, light blue, and green open rectangles indicate the predicted solubility in MeOD relative to the signal intensity in the HR-MAS analysis.

chemical shifts, peak assignment using a combination of several spectral measurements is not necessarily required. The current study describes a technique for statistical and structural metabolic

profiling based on ^1H - ^{13}C HSQC spectra and provides a means of evaluating the extraction process to improve the utilization of plant materials.

CONCLUSIONS

Metabolite profiles of ^{13}C -labeled *A. thaliana* were examined by a combined solution-state and HR-MAS NMR approach during a series of repeated extraction processes. The combination of HR-MAS techniques enabled plant metabolites to be analyzed without loss through several extraction steps. Relationships were revealed between the structural properties and behavior of polar and semipolar metabolites in *A. thaliana*. It was also shown that the distribution of metabolites in a given extraction was influenced not only by the solubility or polarity of the metabolite, but also by the composition of *all* metabolites in the cells. These results will improve our knowledge of *Arabidopsis* polar and semipolar metabolite extraction and the interpretation of the obtained metabolite profiles. We should pay attention to these metabolite behaviors, because the increase or decrease in a particular metabolite may not depend strictly on the plant environment, stimuli, or genotype, but may

be the result of competition among metabolites during the extraction process.

ACKNOWLEDGMENT

This work was supported in part by grants from CREST, the Japan Science and Technology Agency (J.K.), Research and Development Program for New Bioindustry Initiatives of the Bio-oriented Technology Research Advancement Institution (BRAIN; J.K.), and New Energy and Industrial Technology Development Organization (NEDO; J.K.).

SUPPORTING INFORMATION AVAILABLE

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

Received for review August 24, 2009. Accepted January 4, 2010.

AC9019076