Monolithic Silica-Based Capillary Reversed-Phase Liquid Chromatography/Electrospray Mass Spectrometry for Plant Metabolomics

Vladimir V. Tolstikov,[†] Arjen Lommen,[‡] Kazuki Nakanishi,[§] Nobuo Tanaka,[||] and Oliver Fiehn*[,†]

Department Lothar Willmitzer, Max Planck Institute of Molecular Plant Physiology, 14424 Potsdam, Germany, RIKILT (Institute of Food Safety), P.O. Box 230, 6700 AE Wageningen, The Netherlands and Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands, Department of Material Chemistry, Kyoto University, PRESTO JST, Sakyo-ku, Kyoto 606-8501, Japan, and Department of Polymer Science and Engineering, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan

Application of C18 monolithic silica capillary columns in HPLC coupled to ion trap mass spectrometry detection was studied for probing the metabolome of the model plant *Arabidopsis thaliana*. It could be shown that the use of a long capillary column is an easy and effective approach to reduce ionization suppression by enhanced chromatographic resolution. Several hundred peaks could be detected using a 90-cm capillary column for LC separation and a noise reduction and automatic peak alignment software, which outperformed manual inspection or commercially available mass spectral deconvolution software.

Current approaches in functional genomics follow multiple strategies to describe the link between changes in expression of a particular gene and alteration of overall characteristics of the organism, the phenotype. Besides multiparallel analysis of mRNA or proteins, the unbiased relative quantification of all metabolites in a biological system (termed metabolomics[1]) has been favored since biologically, changes on the metabolite level are next to alterations in phenotypes. Metabolomic data provide a signature of the physiological state and reflect specific biochemical processes that were altered between mutant and wild-type lines or strains. Metabolic signatures can be acquired by direct infusion mass spectrometry without chromatographic separation, if only mapping the distances of various mutants or wild-type strains is desired.[2,3] This strategy has been coined "metabolic fingerprinting". However, if understanding the role of genes is aimed at precise biochemical descriptions, each metabolite must be detected and quantified individually.

In plant biology and plant functional genomics, many academic and industrial research groups use the small dicotyledonous weed *Arabidopsis thaliana* as model to solve fundamental and applied research objectives since the complete genome sequence and functional gene annotations are publicly available.[4] It is assumed that the number of metabolites of a specific plant species may easily reach some 5000 individual low molecular weight compounds. In a first proof-of-concept study to tackle the Arabidopsis metabolome, over 300 metabolites were detected by using mature GC/MS technology.[5] Although it has been shown that the number of detected peaks in typical GC/MS plant chromatograms[6] can be multiplied by deconvolution algorithms,[7] the de novo identification of GC/MS peaks remains cumbersome.[8] Therefore, complementary strategies have been considered. First, a comprehensive approach was elaborated by applying different methods on capillary electrophoresis (CE) coupled to mass spectrometry.[9−11] If a single sample was analyzed by three different CE/MS methods, up to 1500 peaks were detectable from *Bacillus subtilis* extracts, including many important primary metabolites. However, CE/MS coupling is far from routine in most laboratories, and it seems to be less suitable for secondary than for primary metabolites. It is therefore desirable to develop LC/MS-based methods that can potentially provide greater flexibility and high peak capacity, including the scale-up for off-line fractionation and peak identification. Polar and water-soluble components can be separated and quantified by replacing classical normal-phase chromatography with hydrophilic interaction chromatography.[12−14]

* Corresponding author: Phone: +49-331-5678216. Fax: +49-331-5678250. E-mail: fiehn@mpimp-golm.mpg.de.

[†] Max Planck Institute of Molecular Plant Physiology.

[‡] RIKILT (Institute of Food Safety) and Plant Research International.

[§] Kyoto University.

[||] Kyoto Institute of Technology.

(1) Fiehn, O. *Plant Mol. Biol.* **2002**, *48*, 155−171.
(2) Allen, J.; Davey, H. M.; Broadhurst, D.; Heald, J. K.; Rowland, J. J.; Oliver, S. G.; Kell, D. B. *Nat. Biotechnol.* **2003**, *21*, 692−695.
(3) Vaidyanathan, S.; Rowland, J. J.; Kell, D. B.; Goodacre, R. *Anal. Chem.* **2001**, *73*, 4134−4144.
(4) The Arabidopsis Genome Initiative. *Nature* **2000**, *408*, 796−815.
(5) Fiehn, O.; Kopka, J.; Dörmann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L. *Nat. Biotechnol.* **2000**, *18*, 1157−1161.
(6) Fiehn, O. *Phytochemistry* **2003**, *62*, 875−886.
(7) Stein, S. E. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 770−781.
(8) Fiehn, O.; Kopka, J.; Trethewey, R. N.; Willmitzer, L. *Anal. Chem.* **2000**, *72*, 3573−3580.
(9) Soga, T.; Ueno, Y.; Naraoka, H.; Ohashi, Y.; Tomita, M.; Nishioka, T. *Anal. Chem.* **2002**, *74*, 2233−2239.
(10) Soga, T.; Heiger, D. N. *Anal. Chem.* **2000**, *72*, 1236−1241.
(11) Soga, T.; Ueno, Y.; Naraoka, H.; Matsuda, K.; Tomita, M.; Nishioka, T. *Anal. Chem.* **2002**, *74*, 6224−6229.
(12) Tolstikov, V. V.; Fiehn, O. *Anal. Biochem.* **2002**, *301*, 298−307.
(13) Schlichtherle-Cerny, H.; Affolter, M.; Cerny, C. *Anal. Chem.* **2003**, *75*, 2349−2354.
(14) Que, A. H.; Novotny, M. V. *Anal. Chem.* **2002**, *74*, 5184−5191.

10.1021/ac034716z CCC: $27.50
Published on Web 00/00/0000

This strategy is tailored for highly polar compounds, and it also allowed separation of medium polar compounds that are normally submitted to reversed-phase chromatography. However, for samples with higher complexity of lipophilic components, use of reversed-phase columns is advantageous due to better chromatographic resolution. Most work has focused on optimization of LC/MS conditions for quantitation of specific compound classes such as saponins,[15] lipids,[16] flavonoids,[17,18] or carotenoids.[19] The idea of metabolomics is now to compromise between the optimal conditions for each class and in doing so to cover as many metabolites and different compound classes as possible in a single chromatographic run. Generally, this needs improved chromatographic separation in addition to more sophisticated MS experiments and better software for unbiased peak finding and quantitation. Two fundamental reasons underpin the need for high chromatographic resolution and peak capacity before mass spectrometric detection and quantitation: first, the existence of multiple isomers for many biologically relevant compounds and, second, ion suppression and adduct formation in the electrospray ionization process.[20] One option to achieve improved chromatographic resolution is by using monolithic silica capillary columns.[21−23] The permeability is up to 30 times greater than a column packed with 5-μm particles, owing to the large throughpores and the much higher porosity of the monolithic structure than a packed bed of particles. Long capillary columns of up to 100 cm long can be easily prepared to provide high column efficiencies. The high permeability allows high-speed operation of the long capillary columns to generate considerably higher column efficiency than a particle-packed column, 50 000−100 000 plates per column, at a modest pressure drop. The use of a high-efficiency column under gradient conditions should be a viable approach to increase peak capacity for small molecules.

In the present paper, we describe an application of monolithic silica C18 capillary columns of 30−90-cm length for the reversed-phase separation of crude plant metabolomic extracts concomitant with full-scan MS detection, followed by unbiased peak area finding and statistical comparison between different samples. Hybrid-type C18 monolithic silica columns were prepared from a mixture of tetramethoxysilane and methyltrimethoxysilane in a fused-silica tubing (0.2-mm i.d., 0.35-mm o.d.), as reported previously,[23] in lengths of 30, 60, and 90 cm. The 90-cm column evaluated by on-column UV detection with a detection window at 80 cm showed 65 000 theoretical plates for alkylbenzenes in 80% methanol at a linear velocity of 1.5 mm/s. The column performance was very similar to that of the previous preparations.[23] The 30- and 60-cm columns are equivalent to this column in terms of plate height. For metabolomics, we have used an ion trap mass spectrometer due to its enhanced sensitivity in full-scan MS compared to quadrupole, triple quadrupole, or QTOF instruments. The system consisted of a Finnigan LCQ DECA mass spectrometer (ThermoFinnigan, San Jose, CA), a Rheos 2000 pump (Flux Instruments AB, Karlskoga, Sweden), and an HTS PAL autosampler (CTC Analytics, Zwingen, Switzerland). It was operated under Xcalibur software (version 1.3, ThermoFinnigan). The LC column was connected to the electrospray interface of the LCQ mass spectrometer without splitting by a modified continuous nanoflow interface obtained from Protana (Odense, Denmark). 2-kV spray voltage was applied using a T-piece connecting the capillary column and a PicoTip silica microflow emitter purchased from New Objective. The temperature of the heated transfer capillary was maintained at 180 °C. Helium collision gas incoming pressure was 2.6 bar, and the ion gauge pressure was 0.89×10^{-5}. Full-scan mass spectra were acquired at $\sim$1 s^{-1} from 100 to 1500 amu at unit mass resolution. For MS^n experiments for compound identification in separate runs, a data-dependent scan was chosen with the wideband activation turned off. The normalized collision energy was set to 40%, and the activation Q was 0.250 with the source fragmentation turned off. Data acquisition has been done in two modes, positive and negative, due to the ability of the LCQ mass spectrometers in providing continuous polarity switching. Analytical liquid chromatography was performed with a splitting T-joint directly after the sample injector using 6.5 mM ammonium acetate (pH 5.5, adjusted by acetic acid) (A) and the highest purity grade acetonitrile available (Biotech grade solvent, 99+%, (SAF), Seelze, Germany) (B) as the mobile phase at ambient temperature. Pump volumes were 50−120 μL/min, and final flow rates of 5−10 μL/min were reached by a precolumn split ratio of $\sim$1:10. ESI and MS settings were optimized for this microflow rate. The linear velocity of mobile phase in column was found to be $\sim$2.6 mm/s for 30- and 60-cm columns and $\sim$1.8 mm/s for a 90-cm column. Separation was accomplished by stepwise gradients from 5% B to 100% B at 57 (30-cm column), 75 (60-cm column), or 110 min (90-cm column) followed by isocratic elution. Final gradient profiles are pasted into the example chromatograms in Figure 1. The sample injection volume was set to 3 μL, which resulted in an actual split injection of $\sim$0.3 μL onto the capillary column.

Figure 1 shows analyses of −15 °C cold methanol extracts of 100 mg fresh weight of ground leaves of *A. thaliana* by monolithic C18 reversed-phase LC/MS. In repetitive injections, average deviation of retention times were found to be ±0.6 min for compounds eluting before 70 min and ±1 min for later-eluting metabolites. Compared to other plant species such as rice (*Oryza sativa*, data not shown), lipophilic metabolic profiles varied largely in abundance and identity of detectable metabolites, which is in accordance with background knowledge of comparative genomics between different species. For example, glucosinolates only occur in Brassicaceae such as Arabidopsis and are absent from all other species. In general, many important lipophilic compound classes can be observed such as glucosinolates, flavonoids, phenolics, anthocyanines, major and minor components of membrane lipids, porphyrins, chlorophylls and their allomers, and miscellaneous and unidentified compounds. In this respect, a good coverage of the plant lipophilic and secondary metabolome is achieved. The

(15) Huhman, D. V.; Sumner, L. W. *Phytochemistry* **2002**, *59*, 347−360.
(16) Welti, R.; Li, W. Q.; Li, M. Y.; Sang, Y. M.; Biesiada, H.; Zhou, H. E.; Rajashekar, C. B.; Williams, T. D.; Wang, X. M. *J. Biol. Chem.* **2002**, *277*, 31994−32002.
(17) Rauha, J. P.; Vuorela, H.; Kostiainen, R. *J. Mass Spectrom.* **2001**, *36*, 1269−1280.
(18) Deng, F. X.; Zito, S. W. *J. Chromatogr., A* **2003**, *986*, 121−127.
(19) Careri, M.; Elviri, L.; Mangia, A. *J. Chromatogr., A* **1999**, *854*, 233−244.
(20) Bahr, U.; Pfenninger, A.; Karas, M.; Stahl, B. *Anal. Chem.* **1997**, *69*, 4530−4535.
(21) Tanaka, N.; Kobayashi, H.; Nakanishi, K.; Minakuchi, H.; Ishizuka, N. *Anal. Chem.* **2001**, *73*, 420A−429A.
(22) Ishizuka, N.; Kobayashi, H.; Minakuchi, H.; Nakanishi, K.; Hirao, K.; Hosoya, K.; Ikegami, T.; Tanaka, N. *J. Chromatogr., A* **2002**, *960*, 85−96.
(23) Motokawa, M.; Kobayashi, H.; Ishizuka, N.; Minakuchi, H.; Nakanishi, K.; Jinnai, H.; Hosoya, K.; Ikegami, T.; Tanaka, N. *J. Chromatogr., A* **2002**, *961*, 53−63.

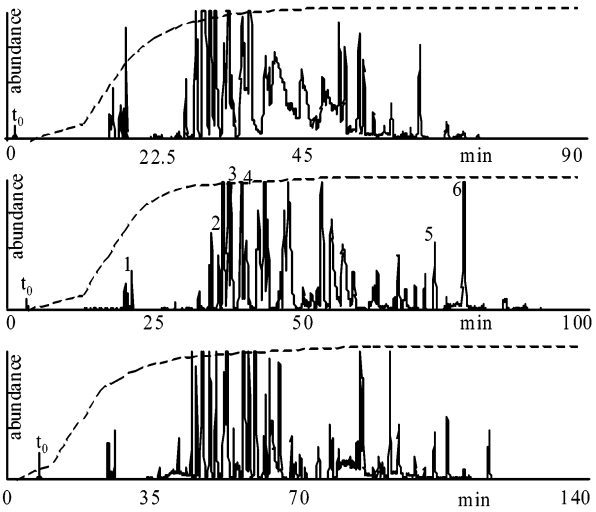

Figure 1. Replicate injections of an Arabidopsis leaf methanol extract on capillary monolithic C18 columns in positive ionization full-scan MS, given as base peak chromatograms. Embedded into the chromatograms (dotted lines) are the actual solvent gradient profiles from the starting conditions of 0−100% B. Upper panel, 0.2 × 300 mm; middle panel, 0.2 × 600 mm; lower panel, 0.2 × 900 mm column. t_0, void volume. Compound class examples detectable in all three chromatograms: 1, flavonoids (e.g., kaempferol); 2, sulfolipids; 3, membrane lipid digalactosyldiacylglycerol; 4, membrane lipid monogalactosyldiacylglycerol; 5, chlorophyll b; 6, pheophytin. Carotenoids elute around 50−60 min and are only detectable by UV-visible under current ESI-MS conditions.

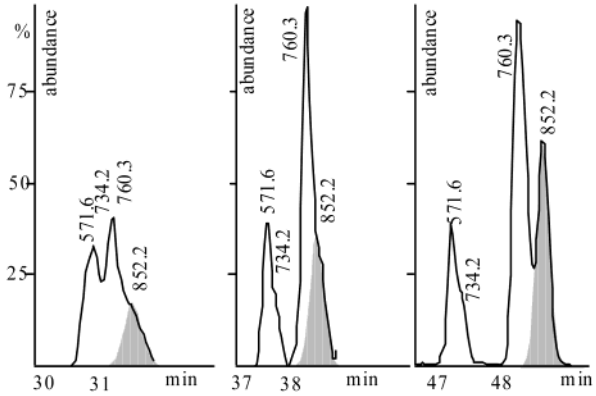

Figure 2. Magnification of three positive ionization full-scan MS chromatograms of replicate injections of a single Arabidopsis leaf extract onto a 200-μm-o.d. capillary monolithic C18 column. Left panel, 300-mm column length; middle panel, 600-mm column length; right panel, 900-mm column length. Gray-shadowed peaks represent the relative peak abundance of the extracted ion trace *m*/*z* 852.2 in comparison to coeluting base peaks *m*/*z* 571.6, 734.2, and 760.3. Relative heights are scaled to the maximum height of 1.41 × 10^7 units.

process of compound identification has been described earlier[12,24] and is mainly based on the fragmentation pathway study, fractionation, off-line accurate mass measurements, and, in special cases, 2D NMR. Under the current electrospray conditions, carotenoids do not produce ions and were only detectable using a LC postcolumn flow splitter and UV−visible detection.

We have compared the benefit of increased resolution using longer monolithic columns. Figure 1 illustrates a comparison of the separation efficiency of an identical sample of Arabidopsis leaf methanol extract injected onto three columns of different lengths (30, 60, and 90 cm) with the same 200-μm i.d. It is clearly indicated that a longer column can provide a much greater number of separated peaks, at least partly based on the reduction of ion suppression. Lower linear velocities or longer gradient times for shorter columns were not so effective as the increase in column length for better resolution.[25] For the example of four coeluting compounds at *m*/*z* 571.6, 734.2, 760.3, and 852.2, it is demonstrated in Figure 2 that not only were the higher theoretical plate numbers helpful for improved chromatographic resolution but, at the same time, the absolute peak areas of these four peaks were changing. This observation is well known among mass spectrometrists and is caused by electrospray ion suppression when several compounds elute at the same time.[26,27] It is seen that better peak separation led to a sudden increase of peak abundance for all four metabolites when comparing the chromatogram run with 600-mm column length to the run with 300-mm column length. A further increase in baseline separation did not result in higher peak areas for the peak pair *m*/*z* 571.6 and 734.2 but did for the peak at *m*/*z* 852.2. These data support the hypothesis that the increase in peak height is due to the better resolution achieved with longer columns, although other effects may also contribute. Since the phenomenon is fundamental to all applications involving a particular ionization source, it is independent from the mass spectrometer that is used. In other words, for highly reliable quantitative metabolomic data, improved chromatographic separation is desired, since the application of hundreds or thousands of stable isotope marker compounds for more accurate quantification seems out of reach.

In terms of metabolome coverage, the total number of detectable peaks is particularly interesting. A careful manual and scan-wise inspection of peaks using the instrument's Excalibur software revealed ~200 distinct components for the Arabidopsis extract using a 900-mm monolith chromatogram, positive/negative ionization switching, and MS/MS triggering. By applying mass spectral deconvolution with Mass Frontier 3.0 software, 172 peaks were found in positive electrospray mode. Unfortunately, MS deconvolution on continuous positive/negative ionization switching is not available with any software to date. This would be highly advantageous since many components are much more efficiently ionized in negative mode than in positive due to dependence of the ionization efficiency on the particular chemical structure of the ionized molecule.

We have tested an α version of the metabAlign software for filtering out statistically significant differences between predefined classes of chromatograms. This program runs the following algorithm: (a) data smoothing by digital filters related to the average peak width, (b) estimation of local noise as a function of retention time and ion trace, (c) baseline correction of ion traces and introduction of a threshold to obtain noise reduction, (d) calculation and storage of peak maximum amplitudes, (e) between-chromatogram alignment using high S/N peaks common to all

(24) Tolstikov, V.; Costisella, B.; Weckwerth, W.; Zhang, B.; Fiehn, O. *Proc. ASMS Conf. Mass Spectrom. Allied Top.* **2002**, *50*, 1−1.
(25) Snyder, L. R.; Glajch, J. L.; Kirkland, J. J. *Practical HPLC Method Development*; Wiley: New York, 1997: Chapter 6, p 188.
(26) Schmidt, A.; Karas, M.; Dulcks, T. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 492−500.
(27) Matuszewski, B. K.; Constanzer, M. L.; Chavez-Eng, C. M. *Anal. Chem.* **2003**, *75*, 3019−3030.

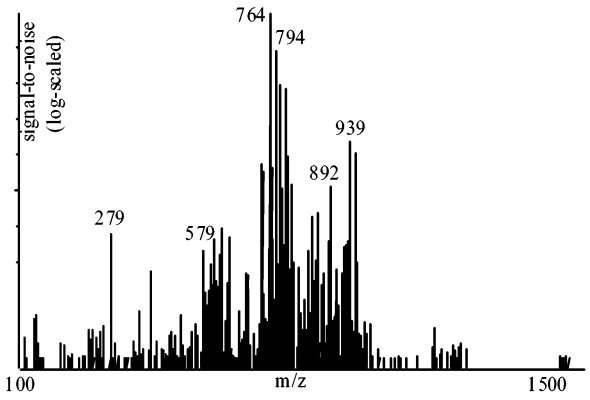

Figure 3. Arabidopsis leaf extract. Distribution of *m*/*z* within the acquisition mass range of 100−1500 amu exceeding S/N > 6. Ions detected for positive ionization full-scan MS chromatogram on a 0.2 × 900 mm monolith C18 column and unbiased peak finding by metAlign software. Intensities are represented in log scale for graphical reasons.

chromatograms ("landmark peaks"), (f) iterative fine alignment by including an increasing number of landmark peaks with lower S/N, and (g) filtering for significant differences at user-defined significance thresholds and minimum *x*-fold ratios.

The same software can be used to get the total number of unique ions present in a particular chromatogram and reducing this number for adduct and fragment ions and corresponding isotope ions. This estimation led to the detection of over 700 different base peaks that exceeded a threshold for S/N > 6 after smoothing. A visualization of the base peak ions found in the chromatogram and its relative signal-to-noise ratios is given in Figure 3. Most of the high-abundant ions were related to metabolite base peaks that were already known to us such as flavonoids or the membrane lipid DGDG. Even if many of the low-abundant peaks might be due to reagent impurities or contamination of sample preparation tubes and devices, the number of real metabolites will still exceed several hundred and thus outperform both MassFrontier and manual inspection with respect to precision and time needed for chromatogram evaluation.

In conclusion, we have briefly described the analysis of very complex metabolite mixtures of plant origin using capillary silica-based monolithic C18 columns coupled to electrospray ion trap mass spectrometry. Compared to shorter monolithic columns, superior resolution, better peak shape, and higher peak capacity have been achieved using a 900-mm monolithic column. We believe that it would be an easily acceptable approach for common LC/MS experimentalists to use long capillary columns based on monolithic silica to get greater peak capacity in LC separations and to use a data-mining software for better use of LC/ESI MS methods in plant metabolomic analyses and related fields such as 1D and 2D separation of tryptic digests in proteomics.

ACKNOWLEDGMENT

We thank Mr. Hiroshi Kobayashi and Hiroshi Kimura for their technical assistance. Capillary monolith columns were manufactured by the Nakanishi and Tanaka groups, and the metabAlign software was developed by Lommen and colleagues at Plant Research International and RIKILT (Institute of Food Safety). This work was financially supported by metanomics GmbH & Co. KGaA, Berlin, Germany, by the Japanese Ministry of Education, Sports, Culture, Science and Technology, and by the project fund "Development of micro HPLC for post-genomic analysis", provided by the Japanese Ministry of Economy, Trade and Industry, through Kansai Bureau of Economy, Trade and Industry, and Osaka Science and Technology Center..

Received for review June 30, 2003. Accepted September 12, 2003.

AC034716Z