the FEBS Journal

MINIREVIEW

Recent applications of NMR spectroscopy in plant metabolomics

Jane L. Ward, John M. Baker and Michael H. Beale

The National Centre for Plant and Microbial Metabolomics, Rothamsted Research, Harpenden, UK

Keywords
fingerprinting; NMR; plant metabolomics; substantial equivalence

Correspondence
M. Beale, The National Centre for Plant and Microbial Metabolomics, Rothamsted Research, West Common, Harpenden, AL5 2JQ, UK
Tel: +44 1582 763133
E-mail: Mike.beale@bbsrc.ac.uk

(Received 19 October 2006, revised 15 November 2006, accepted 20 November 2006)

doi:10.1111/j.1742-4658.2007.05675.x

Recent research has established NMR as a key method for high-throughput comparative analysis of plant extracts. We discuss recent examples of the use of NMR to provide metabolomic data for various applications in plant science and look forward to the key role that NMR will play in data provision for plant systems biology.

Introduction

Nuclear magnetic resonance (NMR) spectroscopy is usually the method of choice for natural product structure determination and it is not surprising that this powerful technique has come to the fore in plant metabolomics. The data requirements for metabolomics are the qualitative and quantitative analyses of the maximum number of metabolites in the highest achievable throughput. Most metabolomics laboratories deploy a range of spectroscopic technologies but use of NMR spectroscopy, particularly as a first pass screen, has a number of advantages over other analytical platforms currently being used. Sample preparation is relatively simple when compared to other analytical methods and a high sample throughput with little instrument drift is readily achieved. NMR is not discriminatory unlike certain mass spectrometry methods that rely on the prior derivatization of metabolites or the ability of them to ionize. Metabolite screening requires maximum sensitivity with a broad compound coverage. For NMR this usually means that only the most sensitive and commonly occurring magnetic nucleus (i.e. ^{1}H) is observed. However, more information on topics such as metabolite flux can be obtained with other nuclei, particularly ^{13}C and ^{15}N. In this minireview we highlight progress in plant metabolomics where ^{1}H-NMR has been used in substantial equivalence and functional genomics studies. In addition, the recent use and prospects for heteronuclear NMR, 2D NMR, liquid chromatography (LC)-NMR and stable isotope labelling experiments are also described.

NMR fingerprinting

Fingerprinting techniques involve collecting spectra of unpurified solvent extracts in standardized conditions and ignore, initially, the problem of making individual

Abbreviations
GM, genetic modification; HSQC, heteronuclear single quantum correlation; J-Res, ^{1}H J-Resolved; PCA, principal component analysis; LC, liquid chromatography; PEG, polyethylene glycol; PLS-DA, prediction to latent structures-discriminant analysis; SPE, solid phase extraction; TCA, tricarboxylic acid.

assignments of peaks in the resulting complex NMR spectra, which contain many overlapping peaks. Multivariate statistical methods such as principal component analysis (PCA) are used to compare sets of spectra to identify clusters of similarity or difference so that conclusions can be drawn about the classification of individual plant samples. The identities of metabolites responsible for differences between classes can be investigated from loadings plots generated by PCA and related techniques. The technique has been recently reviewed by Krishnan *et al*. [1]. Here we discuss recent applications of the technique in a number of key areas.

Substantial equivalence

Regulatory bodies are placing much emphasis on the identification of unintended effects of genetic modification (GM) and there has recently been a drive to establish methods of analysis to screen for these. NMR fingerprinting with multivariate analysis of the data has been used to identify and classify maize seeds, obtained from transgenic plants, into different classes according to changes in metabolites [2]. Prediction to latent structures-discriminant analysis (PLS-DA) methods were used to build a predictive model that could identify GM material by virtue of only 13 variables that were sufficient to explain 90% of the variability in the entire dataset.

In a larger study the 'substantial equivalence' of three transgenic wheats, grown in the field at two different sites for 3 years, has been examined using NMR fingerprinting [3]. Multivariate analysis of the data collected from extracts of flour, milled from the wheat seeds, showed that there was a stronger influence of site and year than there was due to genotype. Although one transgenic line showed elevated levels of carbohydrates, relative to its parent line, most changes detected (e.g., in free amino acids) were due to environment. It was concluded that the plant growth environment has a very significant effect on the metabolome, and that generally differences between control and transgenic wheat lines were within the range of those environmental differences.

Similar NMR fingerprinting studies, assessing the compositional changes occurring in potato tubers after trangenesis, concluded that environmental and cultivar effect were on the whole greater than unintended effects of GM [4]. Approximately 40 GM lines modified in primary carbon metabolism, starch synthesis, glycoprotein processing and polyamine/ethylene metabolism were compared. The most obvious differences revealed by PCA, were between varieties. There were however, significant differences (in proline, trigonelline and other phenolics) between parents and GM lines with modified polyamine metabolism. Generally lines from the other GM groups had altered levels of other compounds relative to the controls, but the differences in mean values amounted only to a two- to three-fold change. It was suggested that, in the context of variability of the whole dataset, such changes did not appear to be important.

Multivariate methods, such as PCA and PLS-DA, build models from the datasets provided and when developing conclusions from these models, the context must be considered. The experimental design is of utmost importance. When dealing with plants, the major differences revealed in principal component 1 (PC1) can sometimes be due to environmental effects. Use of different cultivars can also bias the model. The biological differences (e.g., between GM and non-GM) can be masked and will often only be revealed in the lower PCs. These differences may therefore need to be further studied in refined models where cultivar differences and environmental effects are excluded [3].

Food authenticity and quality control

NMR fingerprinting has been used for many years to authenticate foodstuffs, especially in the beverage industry. A recent study has employed the method to investigate grape quality [5]. The aim was to investigate the effect on grape berry skin metabolites of three cultivars grown over three seasons at five different geographical locations. Using standard methods of NMR data collection, followed by PCA and PLS methods, the predictive modelling was able to pinpoint the spectral areas responsible for a separation according to vintage. No effects due to soil could be discerned and it was concluded that the vintage effect on grape metabolic profiles prevailed over any soil effect.

Quality control issues are becoming increasingly important in the area of phytomedicinal preparations. A recent study on chamomile flowers, employing NMR and chemometrics addressed concerns about variation in composition [6]. The study demonstrated that NMR screening will play an important role in standardization and quality control as legislation is introduced into this area.

Functional genomics

In functional genomics, high-throughput methods that are capable of screening large collections of plants are extremely useful. Metabolomics information can not only assist in a deeper understanding of the complex

interactive nature of plant metabolic networks and their responses to genetic change but also will provide unique insights into the fundamental nature of plant phenotypes in relation to development, physiology and environment. Quantitative metabolite profiling by ^{1}H NMR has recently been used for a genetic study of strawberry fruit quality, a functional study of tomato transformants and a study of *Arabidopsis thaliana* phosphoenolpyruvate transformants [7]. In the tomato study, a comparison of the roots of transformants with wildtypes showed that environmental factors significantly modified the metabolic status of the plants, masking the expression of a given genetic background. Studies on the *Arabidopsis* transformants showed that a decrease in phosphoenolpyruvate carboxylase activity impacted on metabolic profile without compromising the plant growth, supporting previous suggestions that the enzyme had a low influence on the carbon flux through the tricarboxylic acid (TCA) cycle.

Many laboratories are now collecting both metabolomics and transcriptomic datasets from the same tissues and are developing techniques for cross-correlation of these information-rich matrices. No one has yet published such studies with NMR, but a recent example, using GC-MS [8] where the analysis of transcriptome and metabolome datasets gathered in response to sulphur starvation has been modelled, indicates what may be done with such correlative datasets.

Two-dimensional NMR studies

Although 1D NMR studies are extremely useful in classifying similar groups of samples, problems with large numbers of overlapping peaks can make actual identification of large numbers of metabolites difficult. 2D NMR studies can help to overcome these problems. The use of 2D NMR for metabolomics is usually restricted to the characterization of unidentified compounds from the 1D spectra. This can be carried out on the isolated compound (see below). Alternatively, the increased resolution provided by the second dimension can allow for the characterization of components in an unfractionated or partially fractionated mixture. Examples of this include the characterization of tomato juice [9] and the identification of the phenylpropenoids produced by methyl jasmonate treated *Brassica rapa* [10] and *A. thaliana* [11]. The only 2D NMR method that is truly amenable for use as a metabolomic fingerprinting technique is that of ^{1}H J-Resolved (J-Res) NMR, due to its comparatively short acquisition time, relative to other 2D techniques. Figure 1A depicts the complex central region of the conventional ^{1}H-spectrum of a polar extract of *A. thaliana*. The 'skyline projection' [12] (Fig. 1B) generated from the 2D J-Res spectrum (Fig. 1C), is effectively a proton decoupled ^{1}H spectrum. In the projection, multiplets revealed in the 2D J-Resolved plot, are coalesced into single peaks of increased intensity at the chemical shift positions of the multiplet centres. As can be seen in Fig. 1 the result is a spectrum, retaining all the chemical shift and relative intensity data, but with a reduced degree of complexity compared to the conventional ^{1}H spectrum. Choi *et al.* [13] used this technique to investigate the response of tobacco to infection with the mosaic virus. They were able to detect increases in a range of compounds,

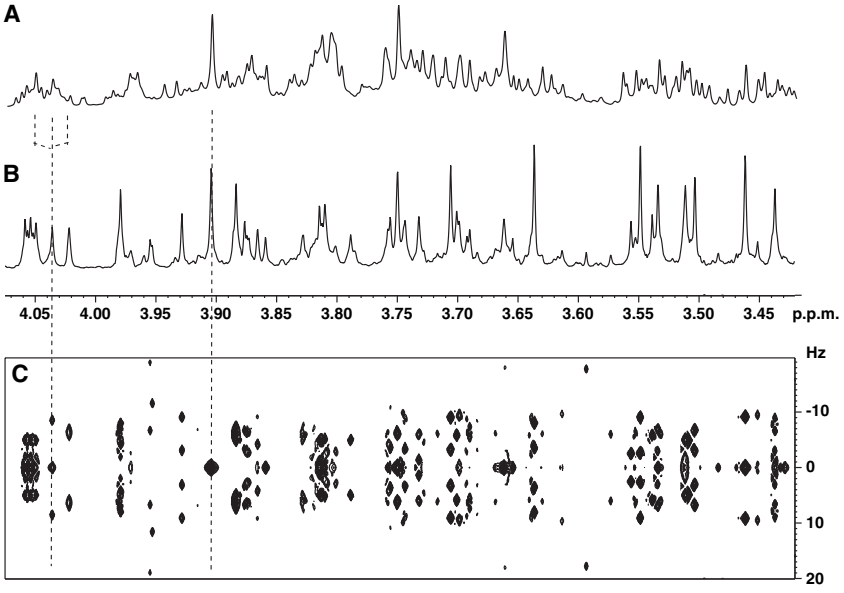

Fig. 1. Carbohydrate region from ^{1}H NMR spectra of an *Arabidopsis thaliana* D_2O-CD_3OD (8 : 2) extract. The conventional 1D spectrum (A); the skyline projection from 2D-J-Res spectroscopy (B); the 2D J-Res spectrum (C). The skyline projection is generated from the 2D plot. A representative triplet at 4.035 p.p.m. (from sucrose) is highlighted and demonstrates the coalescence of a multiplet in the conventional spectrum to yield a single line with a summed area in the skyline projection, whilst singlets (e.g., 3.904 p.p.m.) remain unchanged.

including 5-caffeoylquinic acid, α-linolenic acid analogues, sesquiterpenoids and diterpenoids, and have suggested that these may have a role in systemic acquired resistance.

Metabolite profiling using LC-NMR

Even with the added resolution of 2D NMR techniques, the complete characterization of complex mixtures such as plant extracts by NMR is often impossible. Hyphenating NMR to HPLC alleviates some of the problem by allowing NMR data to be collected on individual components of a mixture. Such on-flow NMR has been used as a means of screening the HPLC profiles of crude lipophilic extracts of aquatic plants for potential algacides [14,15]. The compounds of interest could not be completely identified from the LC-NMR analysis. However, it did give good clues as to the chemical nature of the constituents of the extract, thereby allowing targeted isolation of the most interesting compounds (labdane diterpenes) to be carried out.

The routine use of on-flow LC-NMR for phytochemical analysis is limited by its lack of sensitivity (the previous examples used HPLC runs of > 12 h). The advent of automated solid phase extraction (SPE) peak trapping [16] has removed this problem and allowed LC-NMR to achieve its full potential. The technique has been used to good effect to investigate the composition of an African medicinal plant *Kanahia ianiflora* [17]. Alcoholic extracts of the plant were investigated by analytical scale LC-SPE-NMR using multiple peak trapping, to give sufficient of each of the major peaks to allow their complete characterization using 1D and 2D NMR techniques. Four flavanol-glycosides and three 5α-cardenolides were successfully identified.

Flux analysis by stable isotope tracking

The ability to monitor flux through individual metabolites over time has the potential to offer more to the systems biologist than the single snapshot NMR fingerprinting that is widely used. There have, however, been a growing number of applications in recent years using ^{31}P and stable isotopes such as ^{13}C and ^{15}N to investigate plant metabolism over time and the area has been extensively reviewed [18–20].

Applications using ^{13}C-labelling

Two recent studies serve to illustrate the application of ^{13}C-NMR in flux analysis. Glawischnig *et al.* [21] describe an investigation of carbon flow into starch biosynthesis in maize kernals. Label from added [U-$^{13}C_6$]glucose and [U-$^{13}C_{12}$]sucrose was tracked by ^{13}C NMR measurements on glucose isolated from *de novo* biosynthesized starch, after hydrolysis. ^{13}C tracer studies, using [1,2-$^{13}C_2$]acetate, have also been employed to study the TCA cycle and interacting pathways, in an *in vivo* and *in vitro* metabolomic analysis of rice coleoptiles during anaerbiosis [22]. Peak heights of selected, resolved, ^{13}C NMR signals were normalized against those at time zero and plotted against treatment time to determine the *in vivo* time courses of labelled malate, glutamine, glutamate and γ-aminobutyrate. The study showed that the TCA cycle underwent multiple cycles supporting a separate pool of glutamate, which after decarboxylation yielded γ-aminobutyrate. Diverted carbon was replenished via the glyoxylate cycle reactions. The rice coleoptiles had the ability to reduce the build up of glycolitic by-products (e.g., NADH) by consuming them in various reactions leading to the production of ethanol and amino acids.

^{13}C analysis has also been supplemented by ^{31}P NMR in metabolite profiling studies of perchloric acid extracts of cucumber radicles to reveal changes in phospholipid metabolism in response to osmotic stress and drought tolerance [23]. The radicles were rendered tolerant to desiccation by the addition of polyethylene glycol (PEG). NMR profiling showed increases in sucrose and large decreases in glucose, fructose and the hexose phosphate pool in response to PEG treatment. In addition, three derivatives arising early during phospholipids catabolism appeared in the PEG treated radicles and the study concluded that the metabolic response leading to the re-establishment of drought tolerance was different to that of an osmotic response.

Applications using ^{15}N-labelling

In vivo ^{15}N and ^{31}P NMR studies have also been used to explore symbiotic nitrogen fixation in pea root nodules [25]. The study involved exposing detached pea nodules to $^{15}N_2$ via a perfusion medium, while recording spectra over a time course. After these initial flux measurements, amino acids were extracted and identified by ^{15}N NMR analysis. These studies were complemented using LC-MS. *In vivo* ^{31}P NMR spectroscopy was used to monitor the physiological state of the metabolically active nodules. The investigation also showed (via an unusual ^{15}N chemical shift) that a substantial pool of free ammonium ion was present in active symbiosis. Similar ^{15}N and ^{31}P studies have been employed to investigate primary metabolism in N_2-fixing *Alnus incana-Frankia* symbiotic root nodules [25].

Heteronuclear ^{13}C-^{15}N 2D NMR

The above studies have used stable isotope-feeding studies to answer specific biological problems using a targeted approach. A recent paper demonstrating the utility of stable isotope labelling and 2D heteronuclear NMR for a true metabolomics approach has recently been published [26]. The investigation centred on the metabolic movement of carbon and nitrogen in *A. thaliana*. Ethanol-stress responses were investigated by comparing ^{13}C-labelled wildtype and ^{13}C-labelled ethanol-hypersensitive mutant plants. In a separate study, nitrogen fluxes in ^{15}N-labelled seeds have been analysed during the initiation of germination. Both studies relied on 2D heteronuclear techniques. ^{13}C measurements were made by standard ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC) procedures and the comparisons made by utilizing a spectral subtraction routine. The same principles were applied to the analysis of the ^{15}N nuclei where the authors reported the first ^{1}H-^{15}N HSQC-type NMR experiment to track changes in N-containing metabolites during the onset of germination.

Concluding remarks

NMR has a key role to play in the acquisition of quality assured metabolomic datasets for the systems biologist. It is clear that the now standard ^{1}H-NMR fingerprinting will continue to be the main provider of large datasets for functional and environmental genomics, and in substantial equivalence studies. Hyphenated and 2D techniques play a key role in compound identification and this will lead to more annotation of the 1D fingerprints, as well as answering specific questions by comparative analysis of individual mutants, transgenics or treatments. Systems biology researchers will require metabolomics data that provides information on how plants change over time, whether that be by developmental programming, environmental perturbations or after attack by predators. There is no doubt that NMR fingerprinting and in particular metabolite flux analysis by NMR are now developing to an extent where they will become leading providers of such data.

References

1 Krishnan P, Kruger NJ & Ratcliffe RG (2005) Metabolite fingerprinting and profling in plants using NMR. *J Exp Bot* **56**, 255–265.

2 Manetti C, Bianchetti C, Bizzarri M, Casciani L, Castro C, Ascenzo GD, Delni M, Di Cocco ME, Lagana A, Miccheli A *et al.* (2004) NMR-based metabonomic study of transgenic maize. *Phytochemistry* **65**, 3187–3198.

3 Baker JM, Hawkins ND, Ward JL, Lovegrove A, Napier JA, Shewry PR & Beale MH (2006) A metabolomic study of substantial equivalence of field-grown genetically modified wheat. *Plant Biotechn J* **4**, 381–392.

4 Defernez M, Gunning YM, Parr AJ, Shepherd LVT, Davies HV & Colquhoun IJ (2004) NMR and HPLC-UV profiling of potatoes with genetic modifications to metabolic pathways. *J Agric Food Chem* **52**, 6075–6085.

5 Pereira GE, Gaudillere JP, van Leeuwen C, Hilbert G, Maucourt M, Deborde C, Moing A & Rolin D (2006) H-1 NMR metabolite fingerprints of grape berry: Comparison of vintage and soil effects in Bordeaux grapevine growing areas. *Anal Chim Acta* **563**, 346–352.

6 Wang YL, Tang HR, Nicholson JK, Hylands PJ, Sampson J, Whitcombe I, Stewart CG, Caiger S, Oru I & Holmes E (2004) Metabolomic strategy for the classification and quality control of phytomedicine: a case study of chamomile flower (*Matricaria recutita* L.). *Planta Med* **70**, 250–255.

7 Moing A, Maucourt M, Renaud C, Gaudillere M, Brougisse R, Leboutellier B, Gousset-Dupont A, Vidal J, Granot D, Denoves-Rothan B *et al.* (2004) Quantitative metabolic profiling by 1-dimensional H-1-NMR analyses: application to plant genetics and functional genomics. *Func Plant Biol* **31**, 889–902.

8 Nikiforova VJ, Gakiere B, Kempa S, Adamik M, Willmitzer L, Hesse H & Hoefgen R (2004) Towards dissecting nutrient metabolism in plants: a systems biology case study on sulphur metabolism. *J Exp Bot* **55**, 1861–1870.

9 Sobolev A, Segre A & Lamanna R (2003) Proton high-field NMR study of tomato juice. *Magn Reson Chem* **41**, 237–245.

10 Liang YS, Kim HK, Lefeber AWM, Erklens C, Lefeber AWM, Choi YH & Verpoorte R (2006) Metabolic differentiation of *Arabidopsis* treated with methyl jasmonate using nuclear magnetic resonance spectroscopy. *Plant Sci* **170**, 1118–1124.

11 Hnedrawati O, Yao Q, Kim HK, Linhorst HJM, Erklens C, Choi YH & Verpoorte R (2006) Identification of phenylpropenoids in methyl jasmonate treated *Brassica rapa* leaves using two-dimensional nuclear magnetic resonance spectroscopy. *J Chrom A* **1112**, 148–155.

12 Viant M (2003) Improved methods for the acquisition and interpretation of NMR metabolomic data. *Biochem Biophys Res Comm* **310**, 943–948.

13 Choi YH, Kim HK, Linhorst HJM, Hollander JG, Lefeber AWM, Erklens C, Nuzillard JM & Verpoorte R (2006) NMR metabolomics to revisit the tobacco mosaic virus infection in *Nicotiana tabacum* leaves. *J Nat Prod* **69**, 742–748.

14 Waridel P, Wolfender JL, Lachavanne JB & Hostettmann K (2003) *ent*-Labdene diterpenes from the aquatic plant *Potamogeton pectinatus*. *Phytochemistry* **64**, 1309–1317.
15 Waridel P, Wolfender JL, Lachavanne JB & Hostettmann K (2004) Ent-Labdene glycosides from the aquatic plant *Potamogeton lucens* and analytical evaluation of the lipophilic extract constituents of various *Potamageton* species. *Phytochemistry* **65**, 945–954.
16 Exarchou V, Godejohann M, van Beek TA, Gerothanassis IP & Vervoort J (2003) LC-UV-solid phase extraction-NMR-MS combined with a cryogenic flow probe and its application to the identification of compounds present in Greek oregano. *Anal Chem* **75**, 6288–6294.
17 Clarkson C, Staerk D, Hansen SH & Jaroszewski JW (2005) Hyphenation of solid-phase extraction with liquid chromatography and nuclear magnetic resonance: application of HPLC-DAD-SPE-NMR to identification of constituents of *Kanahia laniflora*. *Anal Chem* **77**, 3547–3553.
18 Ratcliffe RG, Roscher A & Schachar-Hill Y (2001) Plant NMR spectroscopy. *Prog Nucl Magn Reson Spectrosc* **39**, 267–300.
19 Ratcliffe RG & Shachar-Hill Y (2001) Probing plant metabolism with NMR. *Annu Rev Physiol Plant Mol Biol* **52**, 499–526.
20 Ratcliffe RG & Shachar-Hill Y (2005) Revealing metabolic phenotypes in plants: inputs from NMR analysis. *Biol Rev* **80**, 27–43.
21 Glawischnig E, Gierl A, Tomas A, Bacher A & Eisenreich W (2002) Starch biosynthesis and intermediary metabolism in maize kernals. Quantitative analysis of metabolite flux by Nuclear Magnetic Resonance. *Plant Physiol* **130**, 1717–1727.
22 Fan TWM, Lane AN & Higashi RM (2003) *In vivo* and *in vitro* metabolomic analysis of rice coleoptiles revealed unexpected pathways. *Russ J Plant Physiol* **50**, 787–793.
23 Avelange-Macherel M-H, Ly-Vu B, Delaunay J, Richomme P & Leprince O (2006) NMR metabolite profiling analysis reveals changes in phospholipid metabolism associated with the re-establishment of desiccation tolerance upon osmotic stress in germinated radicles of cucumber. *Plant Cell Environ* **29**, 471–482.
24 Scharff AM, Egsgaard H, Hansen PE & Rosendahl L (2003) Exploring symbiotic nitrogen fixation in pea root nodules by *in vivo* ^{15}N Nuclear Magnetic Resonance spectroscopy and liquid chromatography-mass spectrometry. *Plant Physiol* **131**, 367–378.
25 Lundberg P & Lundquist P-O (2004) Primary metabolism in N_2-fixing *Alnus incana-Frankia* symbiotic root nodules studied with ^{15}N and ^{31}P nuclear magnetic resonance spectroscopy. *Planta* **219**, 661–672.
26 Kikuchi J, Shinozaki K & Hirayama T (2004) Stable isotope labelling of *Arabidopsis thaliana* for an NMR-based metabolomics approach. *Plant Cell Physiol* **45**, 1099–1104.