2017

The future of NMR-based metabolomics

John L. Markley  
*University of Wisconsin-Madison, jmarkley@wisc.edu*

Rafael Bruschweiler  
*The Ohio State University*

Arthur S. Edison  
*University of Georgia*

Hamid R. Eghbalnia  
*University of Wisconsin-Madison*

Robert Powers  
*University of Nebraska-Lincoln, rpowers3@unl.edu*

*See next page for additional authors*

Follow this and additional works at: https://digitalcommons.unl.edu/chemfacpub

Part of the [Analytical Chemistry Commons](https://digitalcommons.unl.edu/chemfacpub), [Medicinal-Pharmaceutical Chemistry Commons](https://digitalcommons.unl.edu/chemfacpub), and the [Other Chemistry Commons](https://digitalcommons.unl.edu/chemfacpub)

---

Markley, John L.; Bruschweiler, Rafael; Edison, Arthur S.; Eghbalnia, Hamid R.; Powers, Robert; Raftery, Daniel; and Wishart, David S., "The future of NMR-based metabolomics” (2017). *Faculty Publications -- Chemistry Department*. 127.  
https://digitalcommons.unl.edu/chemfacpub/127

This Article is brought to you for free and open access by the Published Research - Department of Chemistry at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications -- Chemistry Department by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
The future of NMR-based metabolomics

John L Markley¹, Rafael Brüschweiler²,³, Arthur S Edison⁴, Hamid R Eghbalnia¹, Robert Powers⁵, Daniel Raftery⁶ and David S Wishart⁷,⁸

The two leading analytical approaches to metabolomics are mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. Although currently overshadowed by MS in terms of numbers of compounds resolved, NMR spectroscopy offers advantages both on its own and coupled with MS. NMR data are highly reproducible and quantitative over a wide dynamic range and are unmatched for determining structures of unknowns. NMR is adept at tracing metabolic pathways and fluxes using isotope labels. Moreover, NMR is non-destructive and can be utilized in vivo. NMR results have a proven track record of translating in vitro findings to in vivo clinical applications.

Addresses
¹ Biochemistry Department, University of Wisconsin-Madison, 433 Babcock Drive, Madison, WI 53706, USA
² Department of Chemistry & Biochemistry, The Ohio State University, 151 W. Woodruff Ave., Columbus, OH 43210, USA
³ Department of Biological Chemistry & Pharmacology, The Ohio State University, 151 W. Woodruff Ave., Columbus, OH 43210, USA
⁴ Department of Genetics and Biochemistry, Institute of Bioinformatics and Complex Carbohydrate Center, University of Georgia, 315 Riverbend Rd, Athens, GA 30602, USA
⁵ Department of Chemistry, University of Nebraska-Lincoln, 722 Hamilton Hall, Lincoln, NE 68588, USA
⁶ Department of Anesthesiology & Pain Medicine, 850 Republican St, University of Washington, Seattle, WA 98109, USA
⁷ Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada T6G 2E8
⁸ Department of Computing Science, University of Alberta, Edmonton, AB, Canada T6G 2E8

Corresponding author: Markley, John L (markley@wisc.edu)

NMR and its advantages

Despite its lower sensitivity, NMR spectroscopy offers many unparalleled advantages over MS [3**,4**]. NMR
offers a window into observing and rigorously quantifying all of the more abundant compounds present in biological fluids, cell extracts, and tissues without the need for elaborate sample preparation or fractionation. NMR offers advantages for compounds that are difficult to ionize or require derivatization for MS. NMR allows the identification of compounds with identical masses, including those with different isotopomer distributions. NMR is the mainstay for determining structures of unknown compounds. Through the use of stable isotope labels, NMR can be used to elucidate the dynamics and mechanisms of metabolite transformations and to explore the compartmentalization of metabolic pathways. NMR has advantages in drug screening [5]. Finally, site-specific NMR imaging and spectroscopy offer approaches for metabolic studies in living organisms.

Strategies for the identification of metabolites in complex mixtures from NMR data have been reviewed recently [6**]. The most important nuclei in biomolecular NMR studies are ^1^H (proton), ^13^C, ^15^N, and ^31^P. Of these, ^1^H is the most sensitive followed by ^31^P; both are present at near 100% natural abundance. ^31^P NMR is useful for studies of cellular energy states in vivo and ex vivo, but a limitation is that the ^31^P signals from most phosphorylated compounds overlap. One-dimensional (1D) ^1^H NMR is the most widely used NMR approach in metabolomics. Signals are either binned and then analyzed or fitted to patterns of signals corresponding to the metabolites expected to be present in the mixture. The latter approach can be problematic in that many ^1^H signals overlap in ways that offer alternative fitting solutions, a problem that can be overcome by standardizing the analysis in terms of biofluid, solution conditions, data collection protocol, and by employing probabilistic fitting (Bayesil) [7**]. ^13^C NMR signals cover a 200 ppm range compared with 10 ppm for ^1^H and as a consequence are better resolved; however, the low sensitivity of ^13^C (less by a factor of 8 or more) is compounded by its low natural abundance (1.1%). Two dimensional (2D) NMR methods offer improved approaches for unambiguous identification of metabolites in mixtures. These 2D methods include ^1^H–^1^H COSY (correlated spectroscopy), ^1^H–^1^H TOCSY (total correlation spectroscopy), and ^1^H–^13^C HSQC (heteronuclear single-quantum correlation). A widely used software package (rNMR) matches regions of interest in spectra of standards to those in experimental mixtures for compound identifications [8].

Figure 1

Schematic representation of metabolic (blue shaded area).
available for automating metabolite identification from combined TOCSY and HSQC data [9**,10**]. By setting tolerances for the matching of $^1$H and $^{13}$C signals, one can maximize compound identification while minimizing false positives [11]. This approach has been extended to a calculated confidence level for compound identifications from NMR data [12]. Another approach for connecting signals from individual compounds in mixtures is based on searching for statistical correlations among the intensities of NMR signals from various samples [13]. Nuclei present at low natural abundance $^2$H (deuteron), $^{13}$C, and $^{15}$N serve as ideal metabolic tracers [14**].

**Need for standards in NMR metabolomics**

Standard NMR spectra and associated information on small biological molecules are available from freely-accessible databases, including HMDB [2], BMRB [15], TOCCATA [9], and COLMAR [10], but they still cover only a fraction of relevant compounds. A repository has been established for results of metabolomics studies from the NIH Common Fund Centers [16**]. The Coordination of Standards in Metabolomics (COSMOS) Initiative is developing a robust data infrastructure and new data exchange standards (http://nmrml.org) for metabolomics data and metadata to support workflows metabolomics applications [17**]. One of the COSMOS projects is a website (http://metabolomexchange.org) that federates data available from the leading metabolomics data repositories. Best practices and standards have been published for metabolic phenotyping of biological fluids [18**,19]. An open-source platform for complete NMR metabolomics data handling (MVAPACK) has been developed as a step toward establishing best practices for the analysis of metabolic fingerprinting data [20**].

**Sample preparation**

Certain biofluids, for example, cerebrospinal fluid, require little or no preparation for NMR. Others, such as plasma contain proteins and lipids that interfere with NMR spectral quality. Treatment with methanol at solvent-to-serum ratio of 2:1 (v/v) has been shown to remove lipoproteins and minimize the loss of metabolites [21] enabling the detection of about 67 different compounds [22**]. Another promising protocol utilizes the removal of protein by added silica nanoparticles [23**].

**Tagging**

An approach for compounds with overlapping $^1$H signals or present at lower concentration is to tag them with an NMR-active label. Nitrogen-15 with attached hydrogen is an attractive tag because 2D $^1$H-$^{15}$N signals can be acquired at high sensitivity without interference from signals from unlabeled compounds owing to the low natural abundance of $^{15}$N [24,25]. Such tags also provide a permanent positive charge for MS analysis.

**Combining NMR and MS**

As reviewed recently [26**], advances in NMR-based and MS-based metabolomics, including the combination of the two approaches, promise to greatly improve the identification and quantitation of compounds in mixtures. One example is the simultaneous analysis of DI-ESI-MS and 1D $^1$H NMR spectral data to yield accurate mass measurements and class separation scores [27]. Other approaches filter data from one approach against the other to increase the number of compounds confidently identified [28*,29]. Another method identifies compounds by exploiting the principle that abundance/intensity ratios are relatively constant for the same metabolite in different samples [30]. Combined NMR and MS has advantages for isotope tracing experiments and metabolic flux analysis. MS generally quantifies isotopic labeling distributions but even with MS/MS often does not give the specific labeling position, which is available from NMR.

**Quantification**

If 1D $^1$H NMR peaks from a compound are well resolved with acceptable signal-to-noise, their intensities correlate linearly with its relative concentration. To determine absolute concentrations, one adds a standard of known concentration. The cross peak intensities of the same 2D $^1$H-$^{13}$C HSQC spectrum, however, do not correlate linearly with concentration. One can collect spectra of mixtures with known concentrations bracketing those of the unknowns and use these to determine factors that translate peak intensity to concentration [31]. Peak intensities in 2D $^1$H-$^{13}$C HSQC spectra can also be converted to concentration from the slopes generated by spectra utilizing different replicates of the pulse sequence module and the linear extrapolation back to zero time of the peak intensities following the delays from one and two modules yields the ‘HSQC$_{0}$ spectrum’ whose peak intensities are proportional to concentration [32]. Spectral overlaps can be accounted for by methods such as FMLR (fast maximum-likelihood reconstruction) [33]. A new experiment ($^1$H-$^{13}$C QUIPU HSQC) aims to quantify in one map a complex mixture composed of low concentrated metabolites [34]. Another approach, one that requires full $^{13}$C labeling, achieves quantification through the collection of $^{13}$C-$^{15}$C CT-TOCSY spectra and the application of analytical approximations based on the known carbon-backbone topologies [35].

**Applications of metabolomics**

Applications of metabolomics include disease diagnosis, monitoring the effects of medical interventions including drugs, detection of adulteration of food, and analysis of biochemical pathways and their perturbations resulting from mutations, aging, diet, exercise, or life style. A recent study showed how in vivo 1D $^1$H NMR spectroscopy can be used for the simultaneous identification and quantification of co-enzymes that report on cellular function [36**]. Another study used this approach to investigate
alterations in the energy/redox-metabolome in dopaminergic cells exposed to environmental/mitochondrial toxins [37]. Studies of the metabolomics of model organisms are both timely and important for understanding of their different biology [38]. Protocols have been described for studies of the metabolomics of bacteria [39] and plants [40]. Metabolomics, along with activity-guided fractionation followed by structural analysis, constitutes a powerful approach for identifying biologically active compounds for studies in chemical ecology [41]. Metabolomics is used regularly in drug discovery programs to uncover the efficacy, specificity, or toxicity of lead compounds [42]. Metabolomics can provide information on the in vivo mechanism of action and to eliminate compounds likely to cause problems with side effects [43]. Recent studies have utilized metabolomics to search for biomarkers for colon cancer [44] and multiple sclerosis [45].

Future technology

All technologies that increase NMR sensitivity are of extreme importance as are improvements in sample preparation [46]. Approaches to high sensitivity include NMR spectrometers with ultra-high-field magnets operating at 1H resonance frequencies of 1.2 GHz or higher. The first such systems are scheduled for delivery in 2017. Small high-temperature superconducting coils can maximize the signal per sample mass: a 13C-optimized 1.5-mm high temperature superconducting NMR probe has enabled novel 13C NMR studies of natural products [47], and this has been followed up with a 1H–13C dual-optimized NMR probe based on double-tuned high temperature superconducting resonators [48]. These probes take advantage of the excellent peak dispersion of 13C spectra [49], which can be augmented by further 2D 13C–13C experiments, such as INADEQUATE [50].

Hyperpolarization offers an approach for enhanced sensitivity with even higher potential. The underlying physics utilizes the magnetic moment of the unpaired electron, which is roughly 2800 times that of 13C and 6900 times that of 15N, to polarize nuclear spins. First demonstrated by Golman and coworkers [51], studies utilizing hyperpolarized 13C to increase sensitivity are becoming routine. An exciting advance is the discovery of efficient and inexpensive method for hyperpolarizing 15N spins at room temperature. Enhancements are on the order of >10 000 enabling the detection of NMR signals for over an hour [52]. It may become possible to use this approach to tag a range of compounds for metabolomics studies in vivo.

Conclusions

The workshop demonstrated that NMR-based metabolomics promises to continue to play an important role in the studies of complex mixtures of small biological molecules, their metabolic networks, and their interactions with biomacromolecules. Although the development of new and better methods continues to be an integral part, the field needs to focus on developing standardized, enlarged, and integrated databases of NMR data on small molecules as well as archives representing the NMR metabolic fingerprints of standard biological fluids and tissue extracts from humans and model organisms. Standardization of best practices for sample preparation, data collection and analysis should enhance the reproducibility of results within the metabolomics community, while at the same time avoiding the risk of adhering to methods and protocols that are suboptimal in a field that is very much in flux. In order to overcome skepticism in omics [53], it will be advisable for metabolomics to build the ideas of reproducibility and data sharing into every tool and database. Future work is expected to build upon core strengths of NMR spectroscopy, which includes its versatility and specificity in the form of 1D and higher dimensional spectra, its reproducibility, its quantitative ability, its capability for following chemical reactions and flux, its ability to identify compounds and deduce structures of unknowns, and its growing potential for collecting metabolomics data in vivo.

Acknowledgments

The authors thank the Morgridge Institute for Research, Madison WI, for its financial support for the workshop on “NMR-Based Metabolomics”. Research of the authors received support from NIH R01GM109046 and P41GM103399 (to JLM); NIH R01 GM 066041 (to RB); NIH U24 (SECCIM) DK097209-01A1 (to ASE and RB); the Georgia Research Alliance (to ASE); NIH P41GM111135 (to HRE); NIH P20 RR-17675, P30GM103335, R01 CA163649-01A1, and R01 AI087668-01A1 (to RP); NIH R01GM085291 and P30CA15704 (to DR); and Genome Canada support (to DSW). ASE gratefully acknowledges Gary Patti for introducing the term “Metabolic Dark Matter”. JLM thanks Jing Fan for useful comments on the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

● of special interest
● of outstanding interest


4. Nagana Gowda GA, Raferty D: Can NMR solve some significant challenges in metabolomics? J Magn Reson 2015, 260:144-160. The authors present their perspectives on emerging trends in NMR-based metabolomics and NMR’s continuing role in the field with an emphasis on recent and ongoing research from their laboratory.


Review of recent strategies for the identification of metabolites in complex biological mixtures.


The authors present a publicly-accessible system, BAYESIL, for the automated generation of the most probable metabolite profile from 1H NMR spectra of serum or cerebrospinal fluid. BAYESIL offers quantitative NMR spectral profiling with an accuracy meeting or exceeding the performance of trained experts.


The authors introduce the TOCCATA database, which contains complete 1H and 13C chemical shift information on individual spin systems and isomeric states of common metabolites. The database supports identification of compounds present in complex metabolic mixtures at 13C natural abundance from 2D 1H TOCSY and 2D 13C-1H HSQC-TOCSY spectra.


The Complex Mixture Analysis by NMR (COLMAR) 13C-1H HSQC database, which treats slowly exchanging isomers a metabolite as separate species, permits improved detection of compounds present at low concentration. The software is accessible for an interactive web interface.


The authors present an elegant study of the effect of tumor microenvironment on metabolic flux.


The authors describe a carbonyl-selective aminooxy probe incorporating 15N in the aminooxy functional group that specifically reacts with free keto and aldehyde functions, but not carboxylates. This probe enables the exclusive detection of 15N-edited NMR of metabolites that contain a free carbonyl function with differential identification of aminooxy adducts of ketones and aldehydes by their very different chemical shifts. By utilizing 2- and 3-bond, 13C-1H couplings more than 30 carbonyl-containing compounds could be detected by the 15N-edited NMR analysis. As the aminooxy probe contains a permanently charged quaternary ammonium group, the adducts are also optimized for detection by mass spectrometry.


The authors announce an effort sponsored by the Framework Programme 7 EU Initiative ‘coordination of standards in metabolomics’ (COSMOS) to develop a robust data infrastructure and exchange standards for metabolomics data and metadata.


The authors review published studies on NMR-based urine metabolic profiling with the aim of identifying key variables that may affect the results. From this survey, they identify a number of issues that require either standardization or careful accounting in experimental design and provide recommendations for urine collection, sample preparation and data acquisition.


The authors report a newly developed open-source platform (MVAPACK) for complete NMR metabolomics data handling and describe its application to metabolic fingerprinting.


Through investigations of pooled human serum, the authors developed a protocol for removing protein while minimizing the removal of metabolites, which enables the routine identification of 67 blood metabolites including amino acids, organic acids, carbohydrates, and heterocyclic compounds. The proposed protocol, which utilizes 2:1 (v/v) methanol:serum to precipitate protein, was found to be far superior to the previously established acetone/treatment in retaining metabolites.


The authors demonstrate how serum protein can be efficiently removed at physiological pH through attractive interactions with silica nanoparticles. The authors conclude that the combination of nanoparticle treatment and ultrafiltration has a minimal effect on the metabolite content and permits the collection of clean homo/hetero and nonmaclear NMR spectra of the serum metabolome.


The authors describe a carbonyl-selective aminooxy probe incorporating 15N in the aminooxy functional group that specifically reacts with free keto and aldehyde functions, but not carboxylates. This probe enables the exclusive detection of 15N-edited NMR of metabolites that contain a free carbonyl function with differential identification of aminooxy adducts of ketones and aldehydes by their very different chemical shifts. By utilizing 2- and 3-bond, 1H-15N couplings more than 30 carbonyl-containing compounds could be detected by the 15N-edited NMR analysis. As the aminooxy probe contains a permanently charged quaternary ammonium group, the adducts are also optimized for detection by mass spectrometry.


The authors review approaches for combining metabolomics data from mass spectrometry with those from NMR spectroscopy.

The authors present an approach for increasing the number of metabolites detected in urine by combining information from NMR and MS data.


In the SUMMIT MS/NMR approach, the chemical formulas of the mixture components are first identified from accurate masses by MS and then all feasible structures (structural manifold) consistent with these chemical formulas are generated. Next, NMR spectra of each member of the structural manifold are predicted and compared with the experimental NMR spectra in order to identify the molecular structures that match the information obtained from both the MS and NMR techniques.


The authors show how 1H NMR can be used to simultaneously determine levels of key coenzymes: oxidized and reduced forms of nicotinamide adenine dinucleotide (NAD+ and NADH), oxidized and reduced forms of nicotinamide adenine dinucleotide phosphate (NADP+ and NADPH), and adenosine triphosphate (ATP) and its precursors, adenosine diphosphate (ADP) and adenosine monophosphate (AMP).


The authors review strategies for overcoming the inherent low sensitivity of biomolecular NMR spectroscopy.


The authors report a novel NMR probe design simultaneously optimized for 1H and 13C detection. The probe employs novel double-tuned HTS resonators designed for optimal magnetic field homogeneity and minimal electric field that generate strong, uniform, and mutually orthogonal magnetic fields at the 1H and 13C NMR frequencies.


The authors demonstrate how a recently developed method (SABRE-SHEATH) can be used to directly hyperpolarize 15N, magnetization and 15N single spin order, with signal decay time constants of 5.8 and 25.7 seconds.
23 minutes, respectively. The approach generates >10,000-fold enhancements with detectable nuclear MR signals that last for over an hour from molecular tags that can be incorporated into a wide range of biomolecules.


Potential solutions are proposed for the problems of irreproducible scientific studies including improvements in protocols and documentation, consideration of evidence from studies in progress, standardization of research efforts, and scientific workforce development.