Review

Magnetic Resonance in Chemistry

Received: 18 September 2008 Accepted: 7 May 2009 Published online in Wiley Interscience:

(www.interscience.com) DOI 10.1002/mrc.2461

NMR metabolomics and drug discovery

Robert Powers*

NMR is an integral component of the drug discovery process with applications in lead discovery, validation, and optimization. NMR is routinely used for fragment-based ligand affinity screens, high-resolution protein structure determination, and rapid protein-ligand co-structure modeling. Because of this inherent versatility, NMR is currently making significant contributions in the burgeoning area of metabolomics, where NMR is successfully being used to identify biomarkers for various diseases, to analyze drug toxicity and to determine a drug's *in vivo* efficacy and selectivity. This review describes advances in NMR-based metabolomics and discusses some recent applications. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: NMR; metabolomics; drug discovery; disease biomarkers; drug toxicity; principal component analysis; differential NMR metabolomics

Introduction

NMR spectroscopy is playing an integral and continually expanding role in the pharmaceutical industry, especially since high-throughput screening[1–5] and structure-based drug discovery[6–8] have evolved to be the driving forces behind the discovery process for new therapeutics.[9] This process can be divided into three major steps: lead discovery, drug optimization, and clinical validation, and NMR makes invaluable contributions at all stages.[10–15] NMR is the primary analytical tool used to confirm the chemical structure and composition of both synthetic and natural product chemical leads.[16,17] NMR high-throughput ligand affinity screens high-throughput ligand affinity screens (HTS), especially given the growing popularity of fragment-based libraries, are a well-established component of the discovery process.[18,19] NMR HTS are routinely used to both validate and identify novel chemical leads.[20,21] The universal adoption of the fragment-based approach means that the rapid screening of small chemical libraries by NMR enables an exponential growth in the exploration of structural space, well beyond traditional HTS methods.[22–24] In addition to the validation and identification of chemical leads, NMR continues to contribute to lead optimization by determining high-resolution protein solution structures and rapid protein–ligand co-structures.[25,26]

The recent expansion into the analysis of the metabolome has also enabled NMR to contribute to the clinical validation step.[13,27,28] By far, this stage is the most challenging and expensive component of the drug discovery process, where a significant number of failures occur.[29,30] From the analysis of biofluids, tissues, and cell extracts, NMR can measure changes in the metabolome resulting from the biological activity of the drug lead.[31–33] The relative concentration and flux of the hundreds to thousands of small-molecular-weight compounds that comprise the metabolome reflect the state of the system.[34–37] As an illustration, a compound designed to inhibit a specific enzyme will result in changes in the concentration of substrates and products associated with the enzyme's activity. Thus, perturbations in the metabolome result from drug efficacy, selectivity, and toxicity. Additionally, the comparative analysis of the metabolome between healthy and diseased individuals identifies metabolites that can be used as biomarkers for the disease.[38–43]

A major advantage of NMR-based metabolomic studies is the general ease and simplicity of the methodology.[44] In general, biofluids or cell lysates are simply added to a deuterated aqueous buffer to maintain pH and provide a lock signal before transferring to an NMR sample tube to collect a one dimensional (1D) ^{1}H NMR spectrum.[45–47] Because of the inherent variability in biological samples, it is necessary to obtain 10–20 replicates and collect a similar number of NMR spectra so that any observed trends are statistically relevant. This collection of NMR spectra is typically analyzed using an unsupervised statistical technique, such as principal component analysis (PCA).[48,49] PCA reduces the multivariable NMR spectra into the lower dimensional PCA space. Specifically, an NMR spectrum is reduced to a single point in a standard two dimensional (2D) or three dimensional (3D) scores plot. The clustering of NMR spectra in a scores plot determines the relative similarity between the data, where spectra that cluster together indicate a similar metabolome. Accurately interpreting the PCA analysis of NMR spectra requires consistency of sample preparation, data collection, and data processing.[50] It is essential that the observed clustering pattern in the PCA scores plot reflects the anticipated perturbations in the metabolome due to drug activity instead of an artifact from data handling or processing.[51–53] Thus, an additional benefit of NMR-based metabolomics is the minimal sample manipulation, which reduces errors in PCA clustering patterns.

In addition to monitoring global perturbations in the metabolome based on the statistical analysis of NMR spectra, the identity and concentration of the major metabolites affected by the drug are also explored by NMR.[54,55] This enables specific metabolites to be identified as potential disease biomarkers,[38–43] to determine if the drug therapy has toxic side effects,[56,57] and to identify metabolic pathways affected by the drug.[58,59] The ability to rapidly and easily monitor the *in vivo* activity of potential drug candidates at the early stage of drug discovery has significant

* *Correspondence to: Robert Powers, Department of Chemistry University of Nebraska-Lincoln, 722 Hamilton Hall, Lincoln, NE 68588-0304, USA. E-mail: rpowers3@unl.edu*

Department of Chemistry, University of Nebraska-Lincoln, 722 Hamilton Hall, Lincoln, NE 68588-0304, USA

benefits for the effective treatment of human diseases.[60] Clearly, identifying compounds that exhibit diminished *in vivo* activity, poor specificity, or toxicity prior to conducting clinical trials is highly desirable and extremely cost effective.[61] Similarly, using NMR to develop accurate and non-invasive protocols for early disease diagnostics by identifying biomarkers is tremendously beneficial to human health. This review will discuss recent developments and applications of NMR metabolomics that are striving to achieve these goals.

Methodology

Processing of NMR spectra

The obvious appeal of NMR-based metabolomics is the relative ease of the methodology,[28] but the success of the approach requires judicious attention to the uniform preparation of samples and consistent data analysis.[50] Specifically, issues such as long-term storage,[62,63] protein removal,[46] selection of extraction solvent,[45] and tissue preparation[47] can all affect the quality and reliability of the analysis. The advantage of PCA is the extreme sensitivity of the method to subtle spectral differences. This sensitivity may be problematic if the variability in clustering patterns within 2D scores plot result from changes in experimental conditions instead of monitoring perturbations in the metabolome. One such example is the unexpected contribution of NMR spectral noise to PCA clustering.[51] The principal component (PC) analysis of ideal metabolomics data consisting of two NMR samples containing either adenosine 5′ triphosphate (ATP) or an ATP–glucose mixture is shown in Fig. 1. The 10 duplicate spectra were obtained by repeatedly collecting an NMR spectrum utilizing a single sample. Surprisingly, a significant amount of dispersion was observed along the PC2 axis despite the essentially identical data. Even more disturbing was the observation that a single spectrum fell outside the 95% confidence level for the PCA model. Eliminating the noise from the NMR spectra resulted in a ~5× tighter clustering pattern and, more importantly, removed the erroneous data point.

Given the inherent sensitivity of PCA to spectral noise, changes in NMR chemical shifts and peak widths due to pH, temperature, or instrument fluctuation may result in undesirable changes in clustering patterns in 2D scores plots[64] A common approach to minimize these problems is the use of binning, where NMR spectra are divided into regions or 'buckets' with widths of 0.01–0.04 ppm.[65,66] The total peak intensity within these buckets is summed, which results in reduced resolution, but variations between spectra are smoothed out. The binning process itself may also induce errors that are caused by the definition of the bin edge. Ideally, the bin edge should correspond to a baseline region, but variations between spectra may cause a bin edge to occur at a peak. This occurs if a simple and constant bin definition is used for each spectrum. Recent techniques use intelligent binning to optimize the bin edge definition, which does not require a constant bin width. De Meyer *et al.*[67] describe an automated binning protocol that uses variable bin widths and a bin quality factor but does not require reference spectra or user-defined parameters (Fig. 2a). The bin quality factor strives to maximize the peak intensity within the bin while minimizing the peak intensity at the bin edges. Noise bins are discarded.

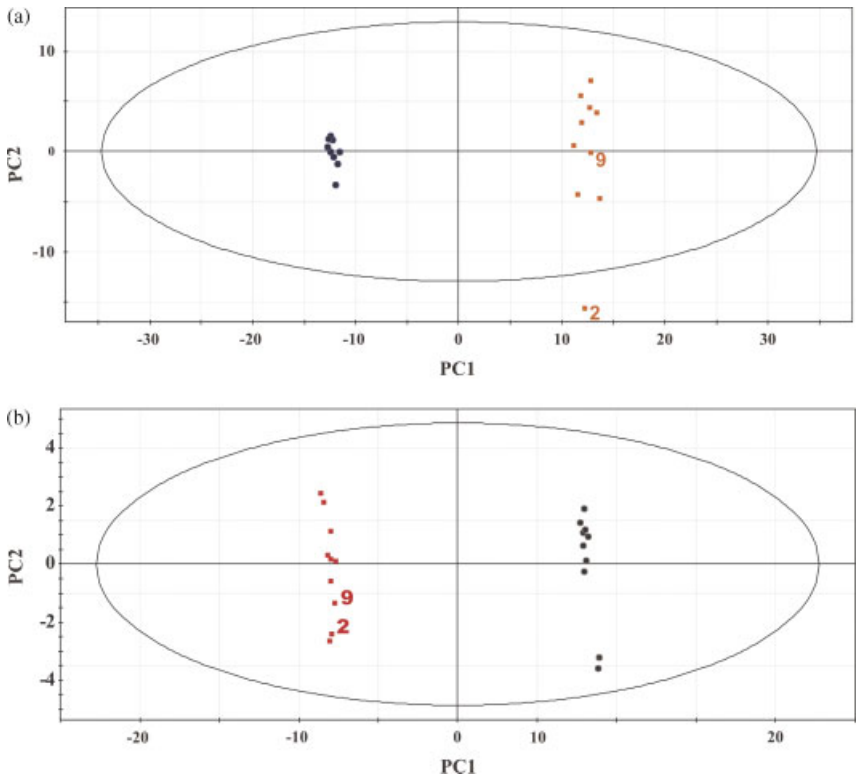


Figure 1. PCA scoring plots of the set of 10 ATP (□) and ATP–glucose (●) NMR spectra (a) with noise and (b) removal of the spectral noise by only binning NMR resonances. The results clearly demonstrate the increased variability and dispersion in the scores plot due to noise (Reprinted with permission from Ref. [51], Copyright 2006 by Elesvier).

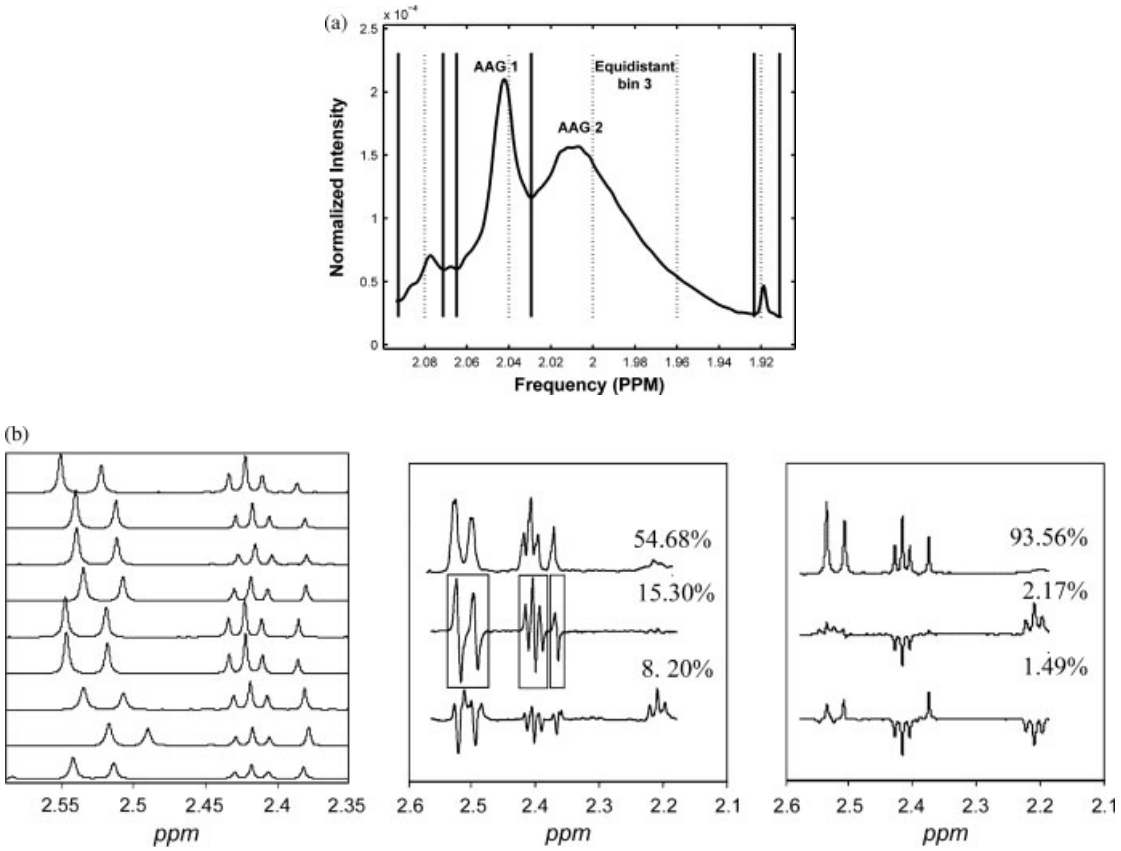


Figure 2. (a) Adaptive Intelligent (AI)-binning clearly isolates the two large R-1 acid glycoprotein peaks (AAG1 and AAG2) in separate bins (solid lines), in large contrast to the bins obtained after standard, equidistant binning (dashed lines), where only equidistant bin 3 is not a mixture of different peaks. (Reprinted with permission from Ref. [67], Copyright 2008 by American Chemical Society). (b) *Left:* the spectral region of interest in nine spectra from the hydrazine dataset, one doublet from the AB type spectrum of citrate, a triplet from 2-oxoglutarate, and the singlet from succinate. *Center:* the first three PCs and their corresponding normalized eigenvalues of the spectral region of interest. *Right:* the spectral region of interest after application of the procedure for the individual peak alignment. (Reprinted with permission from Ref. [68], Copyright 2008 by Elesvier).

Peak alignments between spectra are an alternative approach to binning and assist in the identification and quantification of metabolites present in the metabolome sample.[55,69] A number of approaches have been described to align 1D and 2D ^{1}H NMR spectra collected on metabolomic samples.[68,70–72] Stoyanova *et al*.[68] describe a protocol using PCA to identify regions of a spectrum that experience frequency or phase shift. Specifically, the second PC ($\vec{P}_2$) is sensitive to frequency shifts and will display a derivative shape when frequency shifts are a dominate factor. Subsections of the spectra shown to correlate with $\vec{P}_2$ are then aligned by shifting the frequencies of the peaks to an average frequency. The procedure is repeated until all subsections of the spectrum that correlate with $\vec{P}_2$ are aligned. The analysis of 57 600 ^{1}H NMR spectra of rat urine for a hydrazine toxicity study demonstrates the frequency shift observable in a metabolomics study (Fig. 2b). The PC analysis of the NMR spectra without peak alignment indicates the derivative shape $\vec{P}_2$ to dominate, and frequency shifts contributes ~15% to the total variance. Aligning the NMR spectra and adjusting for the frequency shifts results in a loss of the derivative shape and a drop in contribution to the total variance to ~2%. Thus, the PC analysis is reflecting changes in the metabolome composition (desired outcome) instead of subtle chemical shift changes due to minor changes in pH, salt concentration, and temperature, or instrument stability.

Assigning NMR metabolome spectra

PC analysis of NMR metabolomics data provides a rapid approach to identify global trends and relationships. Alternatively, detailed analysis of the identity and concentration flux of metabolites provides specific comparisons that enable the determination of disease biomarkers and the identification of affected metabolic pathways. This is a relatively challenging endeavor due to the complexity of the metabolome and the lack of reference NMR spectra. First, the metabolome is not completely defined, may contain an infinite number of compounds, and is species dependent,[73] where the number of plant metabolites has been estimated to be 200 000.[74] The Kyoto encyclopaedia of genes and genomes (KEGG) (http://www.genome.ad.jp/kegg/),[75] MetaCyc (http://metacyc.org/),[76] and the Human Metabolome[37] databases contain the extent of what is known regarding metabolic pathways. Similarly, the Madison Metabolomics Consortium database (http://mmcd.nmrfam.wisc.edu/),[36] Human Metabolome database (http://www.hmdb.ca/),[37] and COLMAR Metabolomics Web Server (http://spinportal.magnet.fsu.edu/)[35] are recent efforts to accumulate both ^{1}H and ^{13}C NMR spectra assignments for known metabolites.

These resources are enabling reliable assignments of NMR spectra to determine both the identity and concentration for the majority of metabolites in an NMR sample.[54] Nevertheless, assigning 1D ^{1}H NMR spectra for metabolomic samples is still

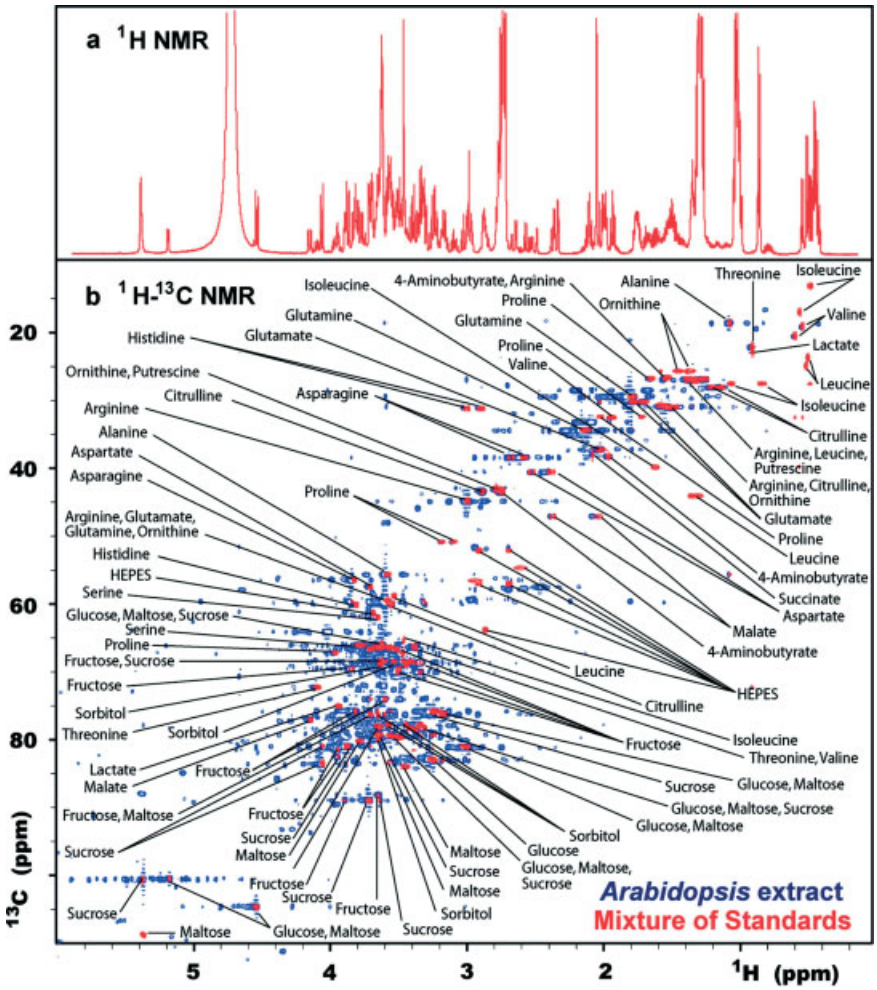
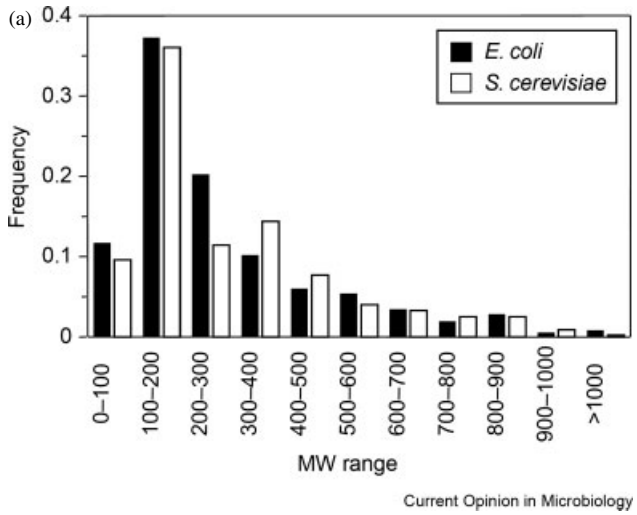


Figure 3. (a) One-dimensional ^{1}H NMR spectrum of an equimolar mixture of the 26 small-molecule standards. (b) Two-dimensional $^{1}H-^{13}C$ HSQC NMR spectra of the same synthetic mixture (red) overlaid onto a spectrum of aqueous whole-plant extract from *A. thaliana* (blue). (Reprinted with permission from Ref. [69], Copyright 2008 by American Chemical Society).

considerably challenging because of significant peak overlap and the presence of uncharacterized metabolites.[77] Instead, the use of 2D NMR techniques is commonly used to analyze the composition of metabolomic samples. The fast metabolite quantification (FMQ) by NMR method described by Lewis *et al*.[69] uses a series of 2D $^{1}H-^{13}C$ heteronuclear single quantum coherence (HSQC) spectra collected for mixtures of standard metabolites over a range of concentrations.[69] An experimental biological sample is then used to collect a 2D $^{1}H-^{13}C$ HSQC spectrum, where peak intensity and chemical shifts are compared against the reference set to identify the metabolites and their corresponding concentration (Fig. 3). Fifty metabolites were identified in the biological extracts from *Arabidopsis*, alfalfa sprouts, and yeast with concentrations ranging from 230 mM to 40 μM.

Integrating NMR and MS metabolomic data

Mass spectroscopy (MS) has traditionally been used to detect perturbations in the metabolome,[78,79] where NMR and MS provide complimentary approaches to the analysis of metabolomic data.[80] An advantage of MS is its relatively high sensitivity and ability to monitor concentration fluxes for minor components that are typically undetected by NMR.[78,79] Conversely, MS typically requires a hybrid approach because of the low-molecular-weight distribution of metabolites (Fig. 4a). Including gas or liquid chromatography to separate compounds with similar molecular weight (MW) may remove or perturb the relative concentration of metabolites. Also, MS is limited to detecting metabolites that are able to ionize well. NMR has similar limitations and is generally restricted to observing metabolites of high concentration. As a result, a number of techniques have been proposed that combine NMR and MS data for the analysis of metabolomic samples.[81–85] A PCA approach that combines 1D ^{1}H NMR data with desorption electrospray ionization mass spectrometry (DESI-MS) data was described by Chen *et al*.[81] The approach was applied to urine samples collected from mice to differentiate between healthy mice and mice with lung cancer (Fig. 4b). Simply, 2D scores plot are calculated separately for the NMR and DESI-MS datasets using a reduced compound dataset. The reduced compound dataset is simply subregions from both the NMR and MS spectra that corresponds to peaks associated with six compounds that distinguish the biological samples. Since the PCs from the NMR 2D scores plots are independent of the DESI-MS data, the NMR PC1 values are simply added to the DESI-MS PC values and become the third dimension in a 3D scores plot. The result is a higher separation of the biological samples in the PC space.



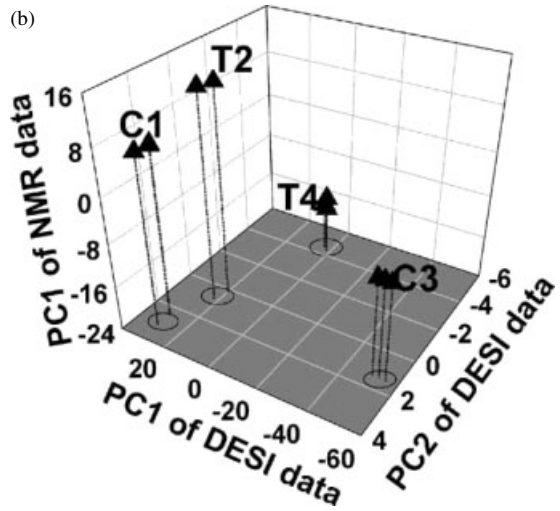


Figure 4. (a) Histogram of molecular weights of typical microbial metabolites (Reprinted with permission from Ref. [73], Copyright 2004 by Elsevier). (b) 3-D score plot combining PCA of NMR and DESI-MS data comparing healthy mice (C1 and C3) with mice with lung cancer (T2 and T4) (Reprinted with permission from Ref. [81], Copyright 2006 by John Wiley & Sons).

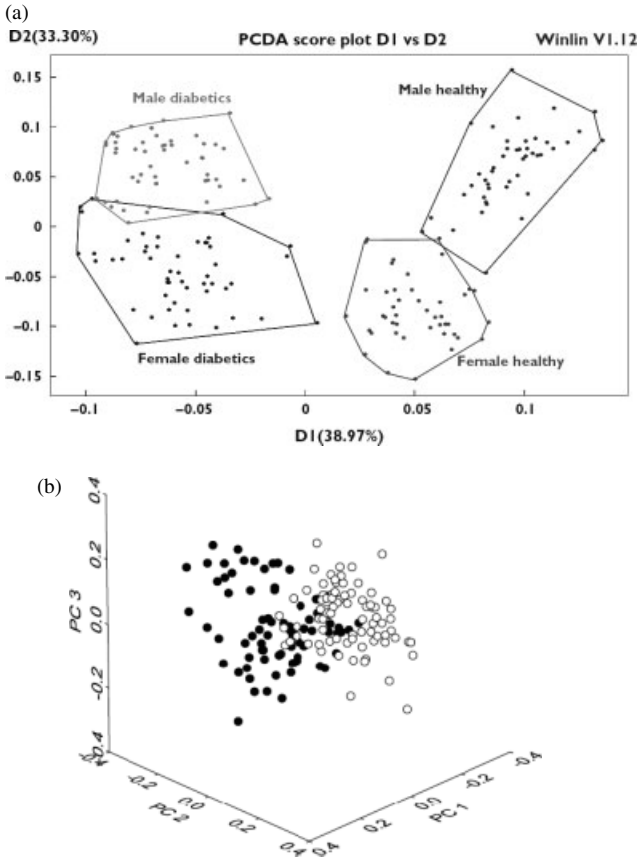


Figure 5. (a) Results of supervised principal component discriminant analysis (PC-DA) of plasma samples from run-in visits (visits 1 and 2). A good distinction between diabetic patients and healthy volunteers as well as separation by gender is observed (Reprinted with permission from Ref. [95], Copyright 2006 by Blackwell Publishing Ltd). (b) 3D PCA scores plot based on covariances of the five NMR bin intensities used in the cancer models. Pancreatic control samples are shown as open circles, while pancreatic cancer samples are shown as solid black circles (Reprinted with permission from Ref. [89], Copyright 2006 by Springer Science).

Applications

Disease biomarkers

One major promise of NMR metabolomics is the identification of biomarkers from biofluids for early disease diagnosis.[28] The approach is straightforward in concept: compare biofluids from healthy and diseased individuals to identify metabolites uniquely correlated with the disease state. Furthermore, it has the added advantage of being rapid and non-invasive, requiring the simple collection of urine, blood, or saliva samples from patients. Of course, there are inherent challenges and limitations in the use of biomarkers.[86] Fundamental variabilities in an individual's metabolome resulting from age, gender, genetics, environmental exposure, behavior, or diet differences may mask the impact of a disease or incorrectly imply the presence of a disease. Other factors, such as the collection, storage, and handling of the biological samples[62,63] or measurement errors,[51,64] may also compromise the correct identification and utility of biomarkers.

NMR-based metabolomics have been used to identify biomarkers associated with a variety of diseases including asthma,[87] arthritis,[88] cancer,[89–91] cardiovascular,[92] diabetes,[93–95] neurodegenerative,[96,97] and pathogen infections.[98] As an illustration, a clinical study described by van Doorn *et al*.[95] demonstrates the use of 1D ^{1}H NMR analysis of blood serum samples to distinguish between healthy volunteers and type 2 diabetes mellitus (T2DM) patients. Eight healthy male and female volunteers and eight male and female patients diagnosed with T2DM had blood serum drawn twice a week over a 6-week period. A PCA of the NMR spectra (Fig. 5a) shows a large differentiation based on both the disease state and gender of the participants in the study. The discrimination is maintained even if the glucose resonances are removed from the NMR spectra. The T2DM biomarkers permitted a further study to determine the effect of thiazolidinedione therapy to treat T2DM. A similar clinical study described by Beger *et al*.[89] was conducted to identify biomarkers for pancreatic cancer, a disease with a high mortality rate (1-year survival rate of 20%) because of difficulties related to early diagnosis. Lipid extracts from plasma samples were collected from 90 healthy volunteers and 100 patients with pancreatic cancer. A subset of only four or five bins from the complete NMR spectra was used to create a partial least squares-discriminant function (PLS-DF) model that statisti-

cally discriminated between healthy individuals and patients with pancreatic cancer (Fig. 5b).

Drug toxicity

Drug toxicity is a very challenging, costly, and pervasive problem in drug discovery,[29,30,99–101] which is primarily caused by the inherent variability in a patient's response to a specific therapy.[102–104] Even in some recent high-profile cases that resulted in the removal of drugs from the market,[105] the vast number of individuals administered the drug did not suffer serious side effects.[106] Generally, only a small percentage of the population suffers serious complications caused by a drug. Ideally, it would be best to identify these individuals prior to starting a drug therapy.[107,108] This would permit general access to the drug and its corresponding benefits to the majority of the population. It is also highly desirable to identify potential drug toxicity events prior to a treatment progressing to serious injury or death.

Similar to its application in identifying biomarkers, NMR metabolomics is becoming an essential tool for the identification and evaluation of drug toxicity.[28,38,56,57] The approach is comparable to the identification of biomarkers: biofluids from animals or patients are analyzed before and after treatment with a drug candidate by 1D ^{1}H NMR and PCA. Any differences in the metabolome that have been associated with serious toxic events, such as liver damage, would be used to identify a toxicity problem with the drug candidate. The approach is demonstrated by a study conducted by Robertson *et al*.,[109] where Wistar rats were treated with two known hepatotoxicants (CCl_4, α-naphthylisothiocyanate) and two known nephrotoxicants (4-aminophenol, 2-bromoethylamine (BEA)). Urine samples were collected daily from the rats and analyzed using 1D ^{1}H NMR and PCA. The 3D scores plot comparing drug treated rats

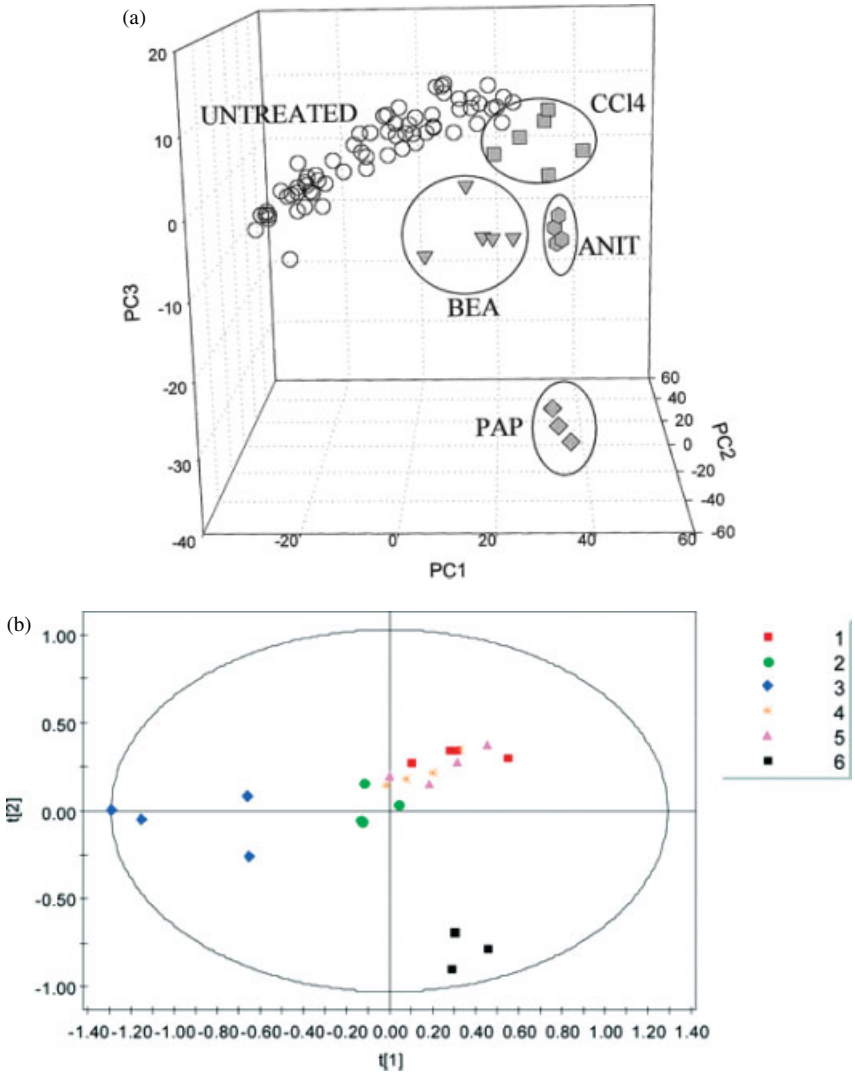


Figure 6. (a) 1D ^{1}H NMR spectra of urine from Wistar rats dosed with two known hepatotoxicants and two known nephrotoxicants. Combined PCA analysis of all toxicant treatments (filled symbols) and all untreated samples (open circles). Toxicant data are as follows: squares = CCl_4, hexagons = α-naphthylisothiocyanate (ANIT), diamonds = 4-aminophenol (PAP), and triangles = 2-bromoethylamine (BEA). The results demonstrate the clear onset of toxicity (Reprinted with permission from Ref. [109], Copyright 2000 by the Society of Toxicology). (b) Score plots from 1D ^{1}H NMR spectra of urine collected from 7-week-old male Han Wistar rats comparing the postdosed samples of the control group (group 1) and the groups dosed with the five compounds (groups 2–6). The samples are colored according to group. Group 3 was determined to have an extreme excretion of choline and the two compounds were excluded as viable drug candidates (Reprinted with permission from Ref. [110], Copyright 2006 by American Chemical Society).

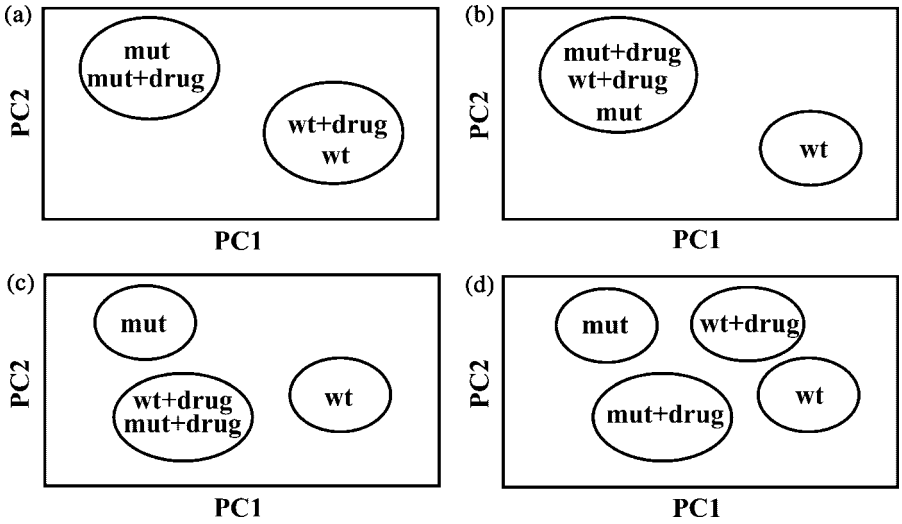


Figure 7. Illustration of the differential NMR metabolomics method. Different clustering patterns in PCA scores plot determine the activity and selectivity of drug candidates. Hypothetical PCA scores plot depict the following scenarios: (a) inactive compound, (b) active and selective inhibitor, (c) active, non-selective inhibition of target and secondary protein, and (d) active, non-selective preferential inhibition of secondary protein. Labels correspond to wild-type cells (wt) and mutant cells (mut) (Reprinted with permission from Ref. [31], Copyright 2006 by American Chemical Society).

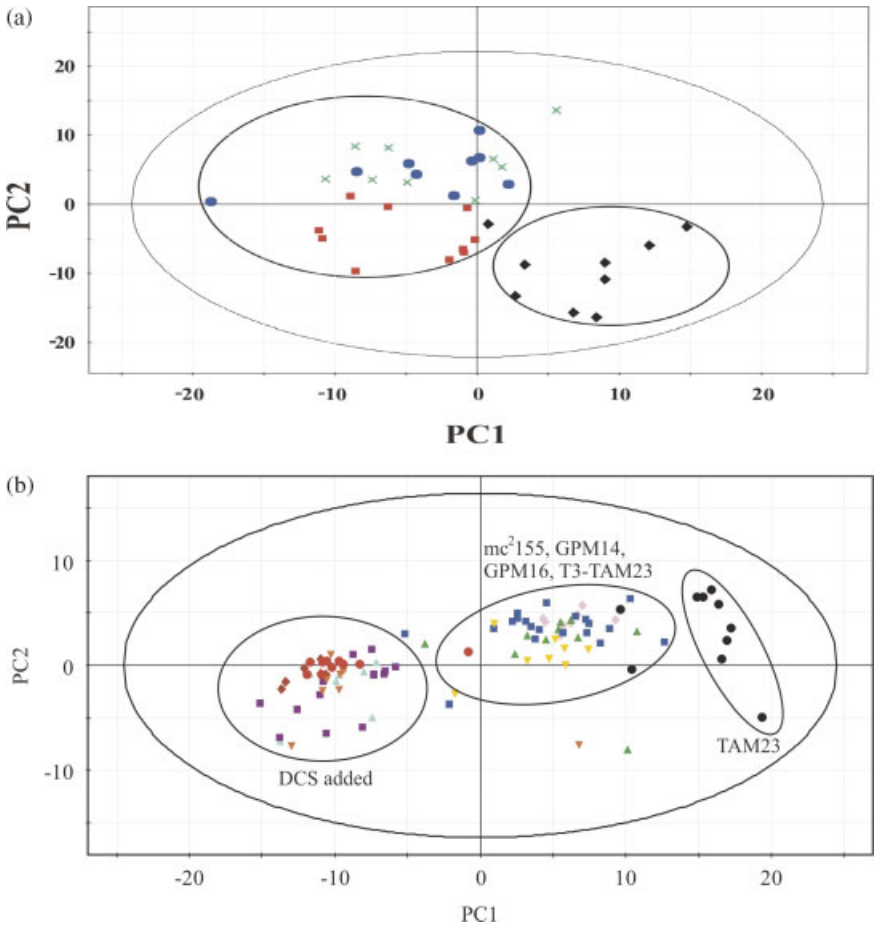


Figure 8. (a) Analysis of the *in vivo* activity of 8-azaxanthine (AZA) in *A. nidulans* targeting urate oxidase. The PCA scores plot comparing *A. nidulans* inactive urate oxidase mutant (*uaZ14*) (×), wild-type with AZA (■), *uaZ14* mutant with AZA (●), and wild-type cells (♦). Results clearly demonstrate the selective activity (see Fig.7b) of AZA (Reprinted with permission from Ref. [31], Copyright 2006 by American Chemical Society). (b) Analysis of the *in vivo* activity of D-cycloserine (DCS) in mycobacteria targeting alanine racemase. PCA scores plot comparing wild-type (mc^2155) (■), inactive D-alanine racemase mutant (TAM23) (●), DCS resistant mutants (GPM14 (♦), GPM16 (▲)), restored D-alanine racemase activity mutant (TAM23 pTAMU3) (▼) mc^2155 with DCS (■), and TAM23 with DCS (●), GPM14 with DCS (♦), GPM16 with DCS (▲), and TAM23 pTAMU3 with DCS (▼). The results clearly demonstrate the active, non-selective inhibition of DCS (see Fig. 7c). The secondary target of DCS is predicted to be D-alanine-D-alanine ligase (Reprinted with permission from Ref. [32], Copyright 2006 by American Chemical Society).

relative to control samples demonstrates a clear differentiation between healthy rats and those with the onset of toxicity (Fig. 6a). Dieterle *et al*.[110] used a similar approach to evaluate potential drug candidates during the early stage of development. Again, Wistar rats were dosed with five different compounds, where urine samples were collected and analyzed by 1D ^{1}H NMR and PCA. A 2D scores plot identified two compounds in Group 3 that were outliers (Fig. 6b). Further NMR analysis determined that the two compounds induced an extreme excretion of choline, which eliminated the two compounds as viable drug leads.

Differential NMR metabolomics

Establishing a drug candidate's *in vivo* efficacy and identifying compounds with toxicity problems at the earliest stages of the drug discovery process is preferable to encountering problems during animal or clinical trials.[111,112] While multiple approaches are being developed to address this challenging endeavor, including *in silco* absorption, distribution, metabolism, and excretion (ADME) protocols,[113–115] the differential NMR metabolomics methodology[31,32] provides a straightforward experimental approach to rapidly ascertain the *in vivo* efficacy and selectivity of drug candidates and to identify potential toxicity issues. The approach compares the metabolome of wild-type (wt) and mutant (mut) cell lines under various environmental stress conditions including drug treatments. In principle, the method can be applied to any type of competent cells (bacteria, fungi, or human tissues). Simply, 1D ^{1}H NMR spectra are collected from a series of cell lysates followed by PCA, where different clustering patterns in the 2D scores plot determines the activity and selectivity of the drug candidate (Fig. 7). In general, four different cell lines corresponding to (i) wt cells, (ii) mutant cells, (iii) wt cells with the drug candidate, and (iv) mutant cells with the drug candidate are grown in ~10–50 ml of culture media. Typically, 10 replicates are prepared for each cell line. The mutant cells have the drug's protein target inactivated. In this manner, if the drug is active and selective, the wt cells in the presence of the drug are expected to cluster with the mutant cells with or without the drug (Fig. 7b). Effectively, the target protein has been inactivated either chemically or genetically resulting in essentially identical changes in the metabolome. Different clustering patterns in the 2D scores plot occur if the drug is inactive, non-selective, or inhibiting an alternative protein (Fig. 7).

Forgue *et al*.[31] demonstrated the general application of the differential NMR metabolomics technique by validating the *in vivo* activity of 8-azaxanthine (AZA), an inhibitor of urate oxidase, in *Aspergillus nidulans*. Small volumes of *A. nidulans* mycelia were prepared in the presence and absence of AZA using wt and a *uaZ14* mutant coding for urate oxidase. As expected, wt cells in the presence of AZA cluster together in the 2D scores plot with the *uaZ14* mutant mycelia with or without the addition of AZA (Fig. 8a). This result suggests AZA is both active and selective *in vivo* (Fig. 7b). A similar study was performed by Halouska *et al*.[32] to determine the *in vivo* target of D-cycloserine (DCS) in mycobacteria. DCS is known to inhibit multiple proteins *in vitro*, but despite being used as a second-tier antibiotic for 50 years, the *in vivo* lethal target is unknown. Both D-alanine racemase (*alr*) and D-alanine-D-alanine ligase (*ddl*) are potential targets, where a *alr* deletion mutant was used to determine whether *alr* is the *in vivo* target of DCS. Again, the differential NMR metabolomics protocol uses 1D ^{1}H NMR spectra from cell lysates of wt and mutant *Mycobacterium smegmatis* in the presence and absence of DCS followed by PCA. Interestingly, all cell cultures in the presence of DCS formed a distinct and separate cluster from the wt and mutant cells (Fig. 8b). This result implies that DCS inhibited a second protein target in *M. smegmatis* (Fig. 7c), where *alr* was not the lethal *in vivo* target of DCS and *ddl* is the likely source of DCS activity in mycobacteria.

Conclusion

The versatility of NMR has enabled the technique to make significant and valuable contributions to the discovery, validation, and optimization of drug leads. In a similar manner, NMR metabolomic methodologies are aiding in the discovery of disease biomarkers, evaluating *in vivo* drug toxicity and assisting in the early evaluation of lead candidates for *in vivo* efficacy and selectivity. The drug discovery process is an extremely challenging and costly endeavor, but these high-risk efforts, when successful, yield tremendous benefits to human health. NMR will continue to expand its important role in drug discovery as evident by the recent advances in NMR-based metabolomics.

References

[1] J. W. Armstrong, *Am. Biotechnol. Lab.* **1999**, *17*, 26.
[2] J. E. Gonzalez, P. A. Negulescu, *Curr. Opin. Biotechnol.* **1998**, *9*, 624.
[3] K. R. Oldenburg, *Annu. Rep. Med. Chem.* **1998**, *33*, 301.
[4] P. B. Fernandes, *Curr. Opin. Chem. Biol.* **1998**, *2*, 597.
[5] B. A. Kenny, M. Bushfield, D. J. Parry-Smith, S. Fogarty, J. M. Treherne, *Prog. Drug Res.* **1998**, *51*, 245.
[6] K. Gubernator, H. J. Boehm, *Methods Princ. Med. Chem.* **1998**, *6*, 15.
[7] H. Kubinyi, *Curr. Opin. Drug Discov. Dev.* **1998**, *1*, 4.
[8] H. Kubinyi, *Die Pharm.* **1995**, *50*, 647.
[9] E. Ratti, D. Trist, *Pure Appl. Chem.* **2001**, *73*, 67.
[10] M. Pellecchia, D. S. Sem, K. Wuthrich, *Nat. Rev. Drug Discov.* **2002**, *1*, 211.
[11] B. J. Stockman, C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *41*, 187.
[12] D. Wishart, *Curr. Pharm. Biotechnol.* **2005**, *6*, 105.
[13] M. Betz, K. Saxena, H. Schwalbe, *Curr. Opin. Cell Biol.* **2006**, *10*, 219.
[14] R. Powers, M. M. Siegel, in *Novel Anticancer Agents: Strategies for Discovery and Clinical Testing*, (Eds: A. A. Adjei, J. K. Buolamwinis), Elsevier Science: New York, **2006**, p 107.
[15] R. Powers, *J. Struct. Funct. Genomics* **2002**, *2*, 113.
[16] E. Fukushi, *Biosci. Biotechnol. Biochem.* **2006**, *70*, 1803.
[17] E. E. Kwan, S. G. Huang, *Eur. J. Org. Chem.* **2008**, *16*, 2671.
[18] P. J. Hajduk, J. Greer, *Nat. Rev. Drug Discov.* **2007**, *6*, 211.
[19] J. Klages, M. Coles, H. Kessler, *Analyst (Cambridge, U.K.)* **2007**, *132*, 693.
[20] H. O. Villar, J. Yan, M. R. Hansen, *Curr. Opin. Cell Biol.* **2004**, *8*, 387.
[21] K. A. Mercier, M. Shortridge, R. Powers, *Comb. Chem. High Throughput Screen.* **2008**, *12*, 285.
[22] K. A. Mercier, K. Germer, R. Powers, *Comb. Chem. High Throughput Screen.* **2006**, *9*, 515.
[23] N. Baurin, F. Aboul-Ela, X. Barril, B. Davis, M. Drysdale, B. Dymock, H. Finch, C. Fromont, C. Richardson, H. Simmonite, R. E. Hubbard, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2157.
[24] A. Schuffenhauer, S. Ruedisser, A. L. Marzinzik, W. Jahnke, M. Blommers, P. Selzer, E. Jacoby, *Curr. Top. Med. Chem. (Sharjah, United Arab Emirates)* **2005**, *5*, 751.
[25] A. E. Ferentz, G. Wagner, *Q. Rev. Biophys.* **2000**, *33*, 29.
[26] J. Stark, R. Powers, *J. Am. Chem. Soc.* **2008**, *130*, 535.
[27] D. G. Robertson, M. D. Reily, J. D. Baker, *Exp. Opin. Drug Metabol. Toxicol.* **2005**, *1*, 363.
[28] J. C. Lindon, E. Holmes, M. E. Bollard, E. G. Stanley, J. K. Nicholson, *Biomarkers* **2004**, *9*, 1.
[29] A. DiMasi Joseph, W. Hansen Ronald, G. Grabowski Henry, *J. Health Econ.* **2003**, *22*, 151.

[30] W. Brains, *Drug Discov. World* **2004**, *5*, 9.
[31] P. Forgue, S. Halouska, M. Werth, K. Xu, S. Harris, R. Powers, *J. Proteome Res.* **2006**, *5*, 1916.
[32] S. Halouska, O. Chacon, R. J. Fenton, D. K. Zinniel, R. G. Barletta, R. Powers, *J. Proteome Res.* **2007**, *6*, 4608.
[33] D. G. Robertson, M. D. Reily, J. D. Baker, *J. Proteome Res.* **2007**, *6*, 526.
[34] V. S. Gomase, S. S. Changbhale, S. A. Patil, K. V. Kale, *Curr. Drug Metab.* **2008**, *9*, 89.
[35] S. L. Robinette, F. Zhang, L. Bruschweiler-Li, R. Bruschweiler, *Anal. Chem.* **2008**, *80*, 3606.
[36] Q. Cui, I. A. Lewis, A. D. Hegeman, M. E. Anderson, J. Li, C. F. Schulte, W. M. Westler, H. R. Eghbalnia, M. R. Sussman, J. L. Markley, *Nat. Biotechnol.* **2008**, *26*, 162.
[37] D. S. Wishart, D. Tzur, C. Knox, R. Eisner, A. C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.-A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D. D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G. E. Duggan, G. D. MacInnis, A. M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B. D. Sykes, H. J. Vogel, L. Querengesser, *Nucleic Acids Res.* **2007**, *35*, D521.
[38] S. Vangala, A. Tonelli, *AAPS J.* **2007**, *9*, E284.
[39] J. P. Shockcor, E. Holmes, *Curr. Top. Med. Chem. (Hilversum, Netherlands)* **2002**, *2*, 35.
[40] G. Navon, H. Burrows, J. S. Cohen, *FEBS Lett.* **1983**, *162*, 320.
[41] J. Pfeuffer, I. Tkac, S. W. Provencher, R. Gruetter, *J. Magn. Reson.* **1999**, *141*, 104.
[42] H. C. Keun, T. M. D. Ebbels, H. Antti, M. E. Bollard, O. Beckonert, E. Holmes, J. C. Lindon, J. K. Nicholson, *Anal. Chim. Acta* **2003**, *490*, 265.
[43] R.-J. A. N. Lamers, J. DeGroot, E. J. Spies-Faber, R. H. Jellema, V. B. Kraus, N. Verzijl, J. M. TeKoppele, G. K. Spijksma, J. T. W. E. Vogels, J. van der Greef, J. H. J. van Nesselrooij, *J. Nutr.* **2003**, *133*, 1776.
[44] J. C. Lindon, J. K. Nicholson, *Trends Anal. Chem.* **2008**, *27*, 194.
[45] H. Wu, A. D. Southam, A. Hines, M. R. Viant, *Anal. Biochem.* **2008**, *372*, 204.
[46] S. Tiziani, A.-H. Emwas, A. Lodi, C. Ludwig, C. M. Bunce, M. R. Viant, U. L. Guenther, *Anal. Biochem.* **2008**, *377*, 16.
[47] C. Y. Lin, H. Wu, R. S. Tjeerdema, M. R. Viant, *Metabolomics* **2007**, *3*, 55.
[48] R. Stoyanova, T. R. Brown, *NMR Biomed.* **2001**, *14*, 271.
[49] J. C. Lindon, E. Holmes, J. K. Nicholson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 1.
[50] B. C. M. Potts, A. J. Deese, G. J. Stevens, M. D. Reily, D. G. Robertson, J. Theiss, *J. Pharm. Biomed. Anal.* **2001**, *26*, 463.
[51] S. Halouska, R. Powers, *J. Magn. Reson.* **2006**, *178*, 88.
[52] M.-E. Dumas, E. C. Maibaum, C. Teague, H. Ueshima, B. Zhou, J. C. Lindon, J. K. Nicholson, J. Stamler, P. Elliott, Q. Chan, E. Holmes, *Anal. Chem.* **2006**, *78*, 2199.
[53] A. C. Kuesel, R. Stoyanova, N. R. Aiken, C.-W. Li, B. S. Szwergold, C. Shaller, T. R. Brown, *NMR Biomed.* **1996**, *9*, 93.
[54] T. W. M. Fan, A. N. Lane, *Prog. Nucl. Magn. Reson. Spectrosc.* **2008**, *52*, 69.
[55] D. S. Wishart, *Trends Anal. Chem.* **2008**, *27*, 228.
[56] J. L. Griffin, M. E. Bollard, *Curr. Drug Metab.* **2004**, *5*, 389.
[57] M. Coen, E. Holmes, J. C. Lindon, J. K. Nicholson, *Chem. Res. Toxicol.* **2008**, *21*, 9.
[58] L. M. Raamsdonk, B. Teusink, D. Broadhurst, N. Zhang, A. Hayes, M. C. Walsh, J. A. Berden, K. M. Brindle, D. B. Kell, J. J. Rowland, H. V. Westerhoff, K. Van Dam, S. G. Oliver, *Nat. Biotechnol.* **2001**, *19*, 45.
[59] J. K. Nicholson, J. Connelly, J. C. Lindon, E. Holmes, *Nat. Rev. Drug Discov.* **2002**, *1*, 153.
[60] D. S. Wishart, *Drugs R&D* **2007**, *8*, 349.
[61] M. Segall, *Eur. Biopharm. Rev.* **2008**, 72.
[62] E. J. Saude, B. D. Sykes, *Metabolomics* **2007**, *3*, 19.
[63] A. D. Maher, S. F. M. Zirah, E. Holmes, J. K. Nicholson, *Anal. Chem.* **2007**, *79*, 5204.
[64] R. A. van den Berg, H. C. J. Hoefsloot, J. A. Westerhuis, A. K. Smilde, M. J. van der Werf, *BMC Genomics* **2006**, *7*, 142.
[65] E. Holmes, P. J. D. Foxall, J. K. Nicholson, G. H. Neild, S. M. Brown, C. R. Beddell, B. C. Sweatman, E. Rahr, J. C. Lindon, M. Spraul, P. Neidig, *Anal. Biochem.* **1994**, *220*, 284.
[66] A. Ross, G. Schlotterbeck, W. Klaus, H. Senn, *J. Biomol. NMR* **2000**, *16*, 139.
[67] T. De Meyer, D. Sinnaeve, B. Van Gasse, E. Tsiporkova, E. R. Rietzschel, M. L. De Buyzere, T. C. Gillebert, S. Bekaert, J. C. Martins, W. Van Criekinge, *Anal. Chem.* **2008**, *80*, 3783.
[68] R. Stoyanova, A. W. Nicholls, J. K. Nicholson, J. C. Lindon, T. R. Brown, *J. Magn. Reson.* **2004**, *170*, 329.
[69] I. A. Lewis, S. C. Schommer, B. Hodis, K. A. Robb, M. Tonelli, W. M. Westler, M. R. Sussman, J. L. Markley, *Anal. Chem.* **2007**, *79*, 9385.
[70] L. Csenki, E. Alm, R. J. O. Torgrip, K. M. Aaberg, L. I. Nord, I. Schuppe-Koistinen, J. Lindberg, *Anal. Bioanal. Chem.* **2007**, *389*, 875.
[71] M. Zheng, P. Lu, Y. Liu, J. Pease, J. Usuka, G. Liao, G. Peltz, *Bioinformatics* **2007**, *23*, 2926.
[72] J. Forshed, I. Schuppe-Koistinen, S. P. Jacobsson, *Anal. Chim. Acta* **2003**, *487*, 189.
[73] D. B. Kell, *Curr. Opin. Microbiol.* **2004**, *7*, 296.
[74] W. Weckwerth, *Annu. Rev. Plant Biol.* **2003**, *54*, 669.
[75] J. Wixon, D. Kell, *Yeast* **2000**, *17*, 48.
[76] R. Caspi, H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer, C. Tissier, T. C. Walk, P. Zhang, P. D. Karp, *Nucleic Acids Res.* **2008**, *36*, D623.
[77] A. Vehtari, V.-P. Makinen, P. Soininen, P. Ingman, S. M. Makela, M. J. Savolainen, M. L. Hannuksela, K. Kaski, M. Ala-Korpela, *BMC Bioinformatics* **2007**, *8*.
[78] I. D. Wilson, R. Plumb, J. Granger, H. Major, R. Williams, E. M. Lenz, *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.* **2005**, *817*, 67.
[79] J. M. Halket, D. Waterman, A. M. Przyborowska, R. K. P. Patel, P. D. Fraser, P. M. Bramley, *J. Exp. Bot.* **2005**, *56*, 219.
[80] Z. Pan, D. Raftery, *Anal. Bioanal. Chem.* **2007**, *387*, 525.
[81] H. Chen, Z. Pan, N. Talaty, D. Raftery, R. G. Cooks, *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1577.
[82] D. J. Crockford, E. Holmes, J. C. Lindon, R. S. Plumb, S. Zirah, S. J. Bruce, P. Rainville, C. L. Stumpf, J. K. Nicholson, *Anal. Chem.* **2006**, *78*, 363.
[83] D. J. Crockford, A. D. Maher, K. R. Ahmadi, A. Barrett, R. S. Plumb, I. D. Wilson, J. K. Nicholson, *Anal. Chem.* **2008**, *80*, 6835.
[84] Z. Pan, H. Gu, N. Talaty, H. Chen, N. Shanaiah, B. E. Hainline, R. G. Cooks, D. Raftery, *Anal. Bioanal. Chem.* **2007**, *387*, 539.
[85] M.-E. Dumas, C. Canlet, L. Debrauwer, P. Martin, A. Paris, *J. Proteome Res.* **2005**, *4*, 1485.
[86] R. Mayeux, *NeuroRx* **2004**, *1*, 182.
[87] S. Carraro, S. Rezzi, F. Reniero, K. Heberger, G. Giordano, S. Zanconato, C. Guillou, E. Baraldi, *Am. J. Respir. Crit. Care Med.* **2007**, *175*, 986.
[88] A. M. Weljie, R. Dowlatabadi, B. J. Miller, H. J. Vogel, F. R. Jirik, *J. Proteome Res.* **2007**, *6*, 3456.
[89] R. D. Beger, L. K. Schnackenberg, R. D. Holland, D. Li, Y. Dragan, *Metabolomics* **2006**, *2*, 125.
[90] K. W. Jordan, L. L. Cheng, *Exp. Rev. Proteomics* **2007**, *4*, 389.
[91] T. M. O'Connell, F. Ardeshirpour, S. A. Asher, J. H. Winnike, X. Yin, J. George, D. C. Guttridge, W. He, A. Wysong, M. S. Willis, M. E. Couch, *Metabolomics* **2008**, *4*, 216.
[92] I. Barba, G. de Leon, E. Martin, A. Cuevas, S. Aguade, J. Candell-Riera, J. A. Barrabes, D. Garcia-Dorado, *Magn. Reson. Chem.* **2008**, *60*, 27.
[93] V.-P. Makinen, P. Soininen, C. Forsblom, M. Parkkonen, P. Ingman, K. Kaski, P.-H. Groop, M. Ala-Korpela, *Mol. Syst. Biol.* **2008**.
[94] R. M. Salek, M. L. Maguire, E. Bentley, D. V. Robtsov, T. Hough, M. Cheeseman, D. Nunez, B. C. Sweatman, J. N. Haselden, R. D. Cox, S. C. Connor, J. L. Griffin, *Physiol. Genomics* **2007**, *29*, 99.
[95] M. van Doorn, J. Vogels, A. Tas, E. J. van Hoogdalem, J. Burggraaf, A. Cohen, J. van der Greef, *Br. J. Clin. Pharmacol.* **2007**, *63*, 562.
[96] A. A. Mahdi, S. Annarao, S. Tripathi, A. Nawab, F. Mahdi, M. Hasan, R. Roy, C. L. Khetrapal, *Open Magn. Reson. J.* **2008**, *1*, 71.
[97] N. Jiang, X. Yan, W. Zhou, Q. Zhang, H. Chen, Y. Zhang, X. Zhang, *J. Proteome Res.* **2008**, *7*, 3678.

[98] X. X. Gao, H. M. Ge, W. F. Zheng, R. X. Tan, *Helicobacter* **2008**, *13*, 103.
[99] M. R. Fielden, K. L. Kolaja, *Exp. Opin. Drug Saf.* **2008**, *7*, 107.
[100] J. L. Stevens, *Chem. Res. Toxicol.* **2006**, *19*, 1393.
[101] I. Kola, J. Landis, *Nat. Rev. Drug Discov.* **2004**, *3*, 711.
[102] S. T. Weiss, H. L. McLeod, D. A. Flockhart, M. E. Dolan, N. L. Benowitz, J. A. Johnson, M. J. Ratain, K. M. Giacomini, *Nat. Rev. Drug Discov.* **2008**, *7*, 568.
[103] S.-F. Zhou, J. Fulcher, *Curr. Pharmacogenomics Person. Med.* **2008**, *6*, 134.
[104] D. L. Mendrick, *Toxicology* **2008**, *245*, 175.
[105] B. Hinz, B. Renner, K. Brune, *Nat. Clin. Pract. Rheumatol.* **2007**, *3*, 552.
[106] M. Meadows, *FDA Consum* **2002**, *36*, 11.
[107] T. A. Clayton, J. C. Lindon, O. Cloarec, H. Antti, C. Charuel, G. Hanton, J.-P. Provost, J.-L. Le Net, D. Baker, R. J. Walley, J. R. Everett, J. K. Nicholson, *Nature* **2006**, *440*, 1073.
[108] J. K. Nicholson, E. Holmes, I. D. Wilson, *Nat. Rev. Microbiol.* **2005**, *3*, 431.
[109] D. G. Robertson, M. D. Reily, R. E. Sigler, D. F. Wells, D. A. Paterson, T. K. Braden, *Toxicol. Sci.* **2000**, *57*, 326.
[110] F. Dieterle, G. Schlotterbeck, A. Ross, U. Niederhauser, H. Senn, *Chem. Res. Toxicol.* **2006**, *19*, 1175.
[111] K. Burley Stephen, F. Park, *Genome Biol.* **2005**, *6*, 330.
[112] K. Sun, J.-H. Pan, *Zhongguo Xinyao Yu Linchuang Zazhi* **2007**, *26*, 783.
[113] B. Singh, A. S. Dhake, D. Sethi, Y. Paul, *Pharma Rev.* **2007**, *5*, 93.
[114] B. Singh, M. Parle, Y. Paul, L. Khurana, *Pharma Rev.* **2007**, *5*, 63.
[115] B. Singh, Y. Paul, M. Grover, D. Sethi, *Pharma Rev.* **2007**, *6*, 179.

