Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics

CrossMark

Darrell D. Marshall, Robert Powers *

Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68588-0304, United States

Edited by David Gadian and David Neuhaus

ABSTRACT

Metabolomics is undergoing tremendous growth and is being employed to solve a diversity of biological problems from environmental issues to the identification of biomarkers for human diseases. Nuclear magnetic resonance (NMR) and mass spectrometry (MS) are the analytical tools that are routinely, but separately, used to obtain metabolomics data sets due to their versatility, accessibility, and unique strengths. NMR requires minimal sample handling without the need for chromatography, is easily quantitative, and provides multiple means of metabolite identification, but is limited to detecting the most abundant metabolites ($\geqslant 1\,\mu M$). Conversely, mass spectrometry has the ability to measure metabolites at very low concentrations (femtomolar to attomolar) and has a higher resolution ($\sim 10^3$–10^4) and dynamic range ($\sim 10^3$–10^4), but quantitation is a challenge and sample complexity may limit metabolite detection because of ion suppression. Consequently, liquid chromatography (LC) or gas chromatography (GC) is commonly employed in conjunction with MS, but this may lead to other sources of error. As a result, NMR and mass spectrometry are highly complementary, and combining the two techniques is likely to improve the overall quality of a study and enhance the coverage of the metabolome. While the majority of metabolomic studies use a single analytical source, there is a growing appreciation of the inherent value of combining NMR and MS for metabolomics. An overview of the current state of utilizing both NMR and MS for metabolomics will be presented.

© 2017 Elsevier B.V. All rights reserved.

Contents

---

* Corresponding author at: University of Nebraska-Lincoln, Department of Chemistry, 722 Hamilton Hall, Lincoln, NE 68588-0304, United States.
   E-mail address: rpowers3@unl.edu (R. Powers).

1. Introduction

Metabolic profiling has existed for hundreds of years and was a common practice in ancient Chinese medicine [1]. At the turn of the century the term "metabolomics"[1] was coined and ignited renewed interest in the field [2–4], which has led to rapid developments and advancements in metabolomics procedures and technologies [2–4]. In effect, metabolic profiling has transitioned from the use of smell and/or taste to the application of advanced spectrometric methods such as nuclear magnetic resonance (NMR) or mass spectrometry (MS) [5–8]. Consequently, metabolomics has become a cornerstone of systems biology [9,10] since it enables unique insights that cannot be obtained from other "omics" methods [10–13]. Metabolomics directly relates a measurable chemical response to a biological event, thereby linking the genotype and phenotype of an organism [14]. Thus, the method is increasingly being used in nutrition science [15], environmental science [16], biomedical research [17] and precision medicine [18]. While interest has expanded exponentially in recent years (Fig. 1), the field of metabolomics is still in an early stage of growth and development, and faces many technical challenges [5,7,8,19,20]. Before it can become a routine tool for the life sciences, many of these current limitations need to be overcome. Similar to other "omics" methods, a single metabolomics study may contain an abundance of data making it difficult to correctly extract the relevant biological information [20–22]. Furthermore, the specific analytical platform chosen for metabolomics will uniquely limit the analysis because of spectral and chemical ambiguity, which, in turn, will yield an imperfect biological picture and incomplete coverage of the metabolome [23]. Sensitivity, resolution, dynamic range, ambiguous assignments, limits of detection, and various other issues are universal problems encountered by *all* analytical methods; there simply is not a single analytical source that will detect the entire metabolome.

Over three decades ago, the fundamental importance and benefit of employing several analytical methods to improve the quality of data analysis and advance compound elucidation was highly touted, which, at the time, produced an alphabet soup of hyphenated methods [24]. Although successful hyphenated techniques rarely expanded beyond combining a form of separation such as liquid chromatograph (LC), gas chromatograph (GC), or capillary electrophoresis (CE) with an analysis tool such as NMR or MS, the need for multiple analytical platforms for complex analysis was readily apparent [25]. This is especially true in the field of metabolomics [23]. Nevertheless, to date, the majority of metabolomics data sets have been acquired using only MS or NMR despite their fundamental complementarity [23]. Fortunately, there is a growing appreciation that combining MS and NMR data greatly improves the coverage of the metabolome and enhances the accuracy of metabolite identification [26–31]. For example, it is well-established that the structure elucidation of an unknown natural product is greatly facilitated by combining the unique information from NMR (*e.g.,* chemical shifts, coupling constants, NOEs, spin systems, *etc.*) with MS (*e.g.,* exact mass [molecular formula], molecular fragments) [32,33].

A metabolomics study is defined by both the capabilities and limitations of the analytical method employed, and may be
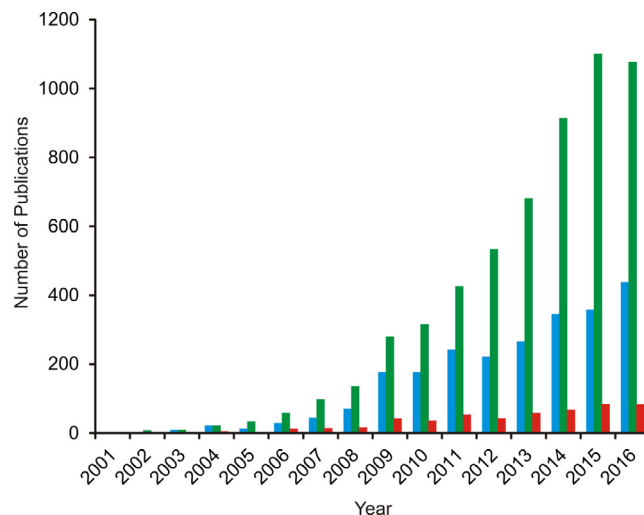

Fig. 1. The chart plots the number of MS metabolomics (green), NMR metabolomics (blue), and combined NMR and MS metabolomics (red) studies published per year from 2001 to 2016. These data were obtained from a keyword search of all documents on PubMed using the key words "MS and metabolomics", "NMR and metabolomics", or "MS and NMR and metabolomics".

significantly hindered if only NMR or MS is used for the analysis. The important advantages of using NMR for metabolomics include a relatively high-throughput, non-destructive data acquisition, minimal sample handling, simple methods for metabolite quantitation, and redundant spectral information to improve the accuracy of metabolite identification [34–36]. Conversely, NMR is limited to detecting only the most abundant metabolites ($\geqslant 1\ \mu M$), while MS has a much higher sensitivity and readily measures concentrations in the femtomolar to attomolar range. MS also boasts higher resolution ($\sim 10^3$–10^4) and dynamic range ($\sim 10^3$–10^4). Conversely, MS only detects metabolites that readily ionize, as a result of which upwards of 40% of chemical libraries are not observable by MS [37,38]. Similarly, ion suppression is a well-known problem in MS, which further reduces the detection of ions of interest due to matrix effects [39]. Simply, the presence of other compounds in the sample or containments from external sources (*e.g.,* plastics, buffers, solid phase, *etc.*) is a 'matrix' that reduces the ability of a specific compound to be ionized through a variety of proposed mechanisms [39]. For example, an ion suppressing agent from the matrix may simply out-compete the compound for available ions. Thus, ion suppression due to matrix effects is a significant concern for metabolomics given the complexity and heterogeneity of metabolomics samples. In effect, the presence of one metabolite may lead to other metabolites being undetected [40]. Thus, MS-based metabolomics typically involves chromatography [41,42] to reduce peak overlap arising from the relatively narrow nominal mass and mass defect distribution of the metabolome [43]. But the use of chromatography may induce biologically irrelevant variations in the metabolome resulting from non-uniform metabolite derivatization, variable metabolite column recovery, metabolite decomposition during derivatization or separation, metabolite ion-suppression due to co-eluting matrix compounds, or misalignment of replicate retention times [44–48]. Simply, NMR and MS have distinct strengths and weaknesses and both uniquely benefit metabolomics.

In recent years, technical advancements in high field magnets, pulse sequences, and cryoprobe technology have led to some significant improvements in the sensitivity and resolution of NMR experiments [23]. This, in turn, has improved the quality of

---

[1] The term metabonomics is often used interchangeably with metabolomics and the primary distinction between the two terms is historical instead of scientific. The word metabolome was first used by Oliver et al. in 1998 to define, similar to the terms proteome or transcriptome, the collection of metabolites present in a cell, tissue or organism. Consequently, metabolomics as defined by Fiehn in 2001 is the 'comprehensive and quantitative analysis of all metabolites'. The term metabonomics was defined in 1999 by Nicholson et al. as 'the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification'.

metabolomics data and has contributed to the observed increase in NMR-based metabolomics studies. Despite these advancements, one-dimensional (1D) ^{1}H NMR spectra are still hindered by a significant amount of peak overlap, even for high field magnets, because of the limited chemical shift dispersion of metabolites [8,49,50]. Thus, these recent improvements in NMR still do not match the sensitivity and resolution advantages of MS [23,51]. Analogous advancements in MS have also occurred, which have led to similar improvements in the quality of metabolomics data sets. High resolution spectrometry, accurate mass, and isotope labeling methods means that virtually every metabolite that is not an isomer produces a unique m/z and should be readily identified [52]. In fact, a majority of metabolomics studies routinely rely on MS (Fig. 1). Additionally, improvements in ionization methods and separation techniques have helped reduce matrix effects due to co-eluting compounds common in complex mixtures, but ion suppression remains a primary issue for MS metabolomics [11,53]. Electrospray ionization mass spectrometry (ESI-MS) sensitivity is directly dependent on a compound's pKa and hydrophobicity, which can be negatively impacted by the heterogeneous composition of a metabolomics sample [54]. For example, if two compounds with different pKa values are present in the same electrospray droplet, only the compound with the lower pKa value may be protonated and detected. Thus, given the various limitations of both NMR and MS, no single analytical platform has the ability to analyze the entire metabolome alone [11,55]. Instead, while some overlap exists, NMR and MS observe a highly complementary set of metabolites [56,57]. Thus, applying both NMR and MS during a metabolomics study will allow for a more comprehensive coverage of the metabolome [57]. The combination of NMR and MS data will also improve the identification of unknown analytes and will increase the accuracy of identifying known metabolites [28,31,51,57–60]. This occurs because metabolite identification will be based on distinct and confirmatory evidence.

A grand challenge in the field of metabolomics is the rapid and accurate identification of metabolites from the variety of complex biological samples routinely analyzed (*e.g.*, tissues, serum and cell extracts) [61,62]. As a consequence, nearly all published metabolomic studies contain at least one misidentified or unidentified metabolite [62]. This is an unfortunate and unavoidable outcome of our limited knowledge of the metabolome (the exact composition is currently unknown), the severe limitations in the software and databases available for metabolite identification, and the routine reliance on a single analytical method. Despite the routine application of NMR and MS to elucidate the structures of natural products for drug discovery [32,33], most metabolic studies still rely on only NMR or MS spectral data for metabolite identification based on database searches [63]. Unfortunately, metabolic databases only contain, at most, a few thousand reference NMR or MS spectra of known metabolites. There is also minimal, if any, coordination between the various metabolomics databases. This leads to a significant amount of redundancy between databases, and also results in unique data being present in individual databases, requiring an investigator to search across multiple databases. Also, many databases are cumbersome and utilize simple search algorithms, which only allow for a single spectral category per query [63]. Finally, most metabolomics databases are limited to either NMR or MS reference data. Again, this is a result of the fact that the metabolomics field has evolved to support only a single analytical source. Thus, it is not possible to easily search metabolomics databases for simultaneous matches against NMR and MS spectral data. As a result, metabolomics will greatly benefit from the creation of a unified NMR and MS database and the merging of data. This is especially true given the growing recognition of the value of combining NMR and MS for metabolite assignments [28,29,51,57,58].

Multivariate statistical methods are normally applied to metabolomics data sets in order to simplify and expedite the data analysis [12,64]. While chemometric techniques regularly streamline the process of data analysis, these advanced multivariate statistical techniques are routinely used incorrectly, lack proper validation, and have, unfortunately, lead to a proliferation of erroneous data in the scientific literature [65,66]. This problem becomes compounded when an investigator is combining multiple analytical sources. In most cases, the samples, data, and analysis are done separately. The NMR and MS data sets are not integrated into a single chemometrics model because, until recently, the field lacked software capable of handling data from multiple analytical sources [57,67,68]. Thus, any observed changes in the NMR and MS spectra are not statistically correlated and, importantly, the metabolites and pathways separately identified by NMR and MS may not be biologically related. Only a minimal number of studies have been reported that actually utilized both NMR and MS in a single chemometrics model [69–72].

As outlined above, NMR and MS are routinely, but separately, used for metabolomics studies despite their complementary strengths. Nevertheless, there is a growing recognition of the benefits of combining the two techniques for metabolomics as evident by the increasing number of published manuscripts that utilized both NMR and MS (Fig. 1). NMR and MS have been combined in a number of ways that includes: (1) physically interfacing NMR and MS hardware, (2) chemical modification of samples by derivatization of metabolites with compounds that display unique characteristics for MS and NMR detection, (3) stable isotope tracing by isotopically labeling metabolites, (4) using combined cheminformatics techniques on MS and NMR data sets for an accurate and rapid analysis, or (5) unique data handling and data mining techniques which correlate trends in both data sets by using multivariate statistical-based methods [29,30,58]. Herein, we provide illustrative examples for each of these methods and present an overview of the various benefits derived from combining NMR and MS for the analysis of metabolomics samples.

2. Brief overview of metabolomics

2.1. Current state of NMR and MS metabolomics

Both NMR- and MS-based metabolomics have been extensively reviewed in the recent scientific literature [3,11,13,23,73–80]. Similarly, a number of reviews describing the proper handling, preparation, and extraction of metabolic samples from urine, serum, cell cultures, tissue cultures, and a variety of other biological sources are readily available [36,81–85]. So, only a very brief overview of the general procedure for metabolic profiling will be presented here, primarily to highlight important differences when NMR and MS are combined for a metabolomics study (Fig. 2) [86]. In principle, metabolomics is a relatively straightforward method: the metabolome is harvested or extracted from two or more groups (*e.g.,* healthy vs. diseased), and an analytical technique (*e.g.,* NMR or MS) is used to acquire a spectral profile of each metabolic sample. Then, multivariate statistical techniques (*e.g.,* principal component analysis [PCA], orthogonal projections to latent structures [OPLS], *etc.*) are used to determine whether the metabolomes differ and, if they do, to identify the spectral features (*i.e.,* the metabolites) defining group separation. A typical outcome of a PCA or OPLS model is a scores plot, where each NMR or MS spectrum has been reduced to a single point in PC-space. The relative clustering of the spectra in the scores plot identifies group membership. Similarly, a backscaled loadings plot is often generated from the OPLS model that is a pseudospectrum of the original NMR or MS data where the relative intensity and direction of the spectral peaks indicates

the contribution and correlation of the peak to the group separation. Alternatively, a heatmap can be generated that compares all of the relative metabolite concentration changes per replicate. Following hierarchal clustering, if all or most of the replicates from the same group cluster together, then the heatmap identifies key metabolite concentration changes that define each group.

Since the numbers of replicates for a metabolomics data set are typically far fewer then the number of variables, overfitting the data, especially for supervised techniques like PLS or OPLS, is a serious concern [65,66,87]. In fact, a PLS/OPLS model can produce the appearance of a clear group separation even for noise or completely random data [88]. As a result, model validation is a critical final step before any biological interpretation of the model is reliable. CV-ANOVA [89], response permutation tests [87], and receiver operating characteristic (ROC) curves [90] are routinely used to provide standard p values and assess model validity. With a validated model, the observed metabolite changes can then be used to generate a metabolic network and identify the important metabolic pathways associated with the phenotype [91].

Unlike the proteome and genome, the metabolome is relatively unstable and is easily perturbed by the handling and processing of the sample. For example, metabolites have different enzyme turnover rates and different temperature stabilities [92–94]. Thus, a key concern is avoiding any changes in the metabolome that may be induced from sample preparation. Therefore, fast and uniform sample preparation protocols, rapid quenching of enzymatic activity, keeping samples cold throughout, and randomizing samples through the entire sample handling and data collection procedure are all important details that require optimization for a successful metabolomics study. Since biological samples are the materials that are analyzed, there is inherently a large natural variance to the data. Thus, to obtain statistical significance a maximal number of replicates, within practical constraints, is highly desirable [36]. Similarly, the data need to be properly normalized, scaled and aligned to account for both biological and instrument variance, and to remove bias due to the large range of metabolite concentrations. Again, various normalization, scaling and alignment algorithms are available and their utility for metabolomics has been previously reviewed [95–97]. Finally, the raw NMR or mass spectral data need to be properly processed before a reliable chemometrics model can be generated. For NMR, processing includes Fourier transformation, phasing, baseline correction, apodization, zero filling and chemical shift referencing. Similarly, MS requires centroiding, de-noising, de-isotoping, deconvolution and peak alignment. It cannot be overstated that the resulting chemometric models are incredibly sensitive to all aspects of the sample handling and data processing. Thus, changing processing details, such as a different apodization function or deconvolution algorithm will likely lead to a different chemometrics model and, potentially, a different biological interpretation. Consequently, great care must be taken in optimizing data processing protocols in order to avoid unintended biases in the data analysis. The complexity of the situation expands exponentially if two or more analytical sources are combined to generate a single chemometrics model simply because of the greater number of possible processing protocol combinations. Thus, while metabolomics is conceptually quite simple and straightforward to conduct, in practice it is difficult to do correctly because of the complex number of steps and choices available, with multiple sources of error encountered at each step of the process.

2.2. Analyte compatibility for analysis by NMR and MS metabolomics

NMR and MS are the most popular analytical methods used for metabolomics because of the wide array of chemical species both techniques can probe and are likely to encounter in a complex and heterogeneous metabolite mixture. MS is generally coupled with a separation technique, with the majority of metabolomics studies utilizing LC-ESI-MS [11,64,98]. The coupling of MS with LC or GC is critical for reducing ion suppression, spectral complexity and spectral overlap. To further reduce ion suppression, samples are typically acidified and salt concentrations are reduced, usually with a desalting column [99]. Although NMR is occasionally paired with a separation method, the hardware for an online LC-NMR system can be somewhat crude or cumbersome, and the process of manually collecting fractions is extremely time-consuming. Moreover, a hybrid LC-NMR system does not provide a substantial benefit since individual metabolites are readily detectable by NMR without chromatographic separation [100].

Besides different chromatographic needs, MS and NMR require different sample conditions. NMR samples are commonly prepared in a buffered deuterated solvent to provide a lock signal and maintain a constant pH of 7.4 in order to minimize chemical shift deviations and replicate physiological conditions [9,99]. In general, detecting a metabolite by NMR is not dependent on the sample condition. However, NMR is a relatively insensitive technique and requires maximizing metabolite concentrations to ensure detection. On average, NMR metabolomics requires a minimal sample volume of approximately 30–600 μL, with metabolite concentrations of approximately μM to mM. Because of the higher sensitivity of MS, the concentration requirements can be significantly reduced to as low as nM to pM, with sample volumes on the order of a few microliters. While MS is intrinsically more sensitive than NMR, detecting a specific metabolite does require the metabolite to be efficiently ionized. Different groups of metabolites will preferentially ionize under drastically different experimental

Fig. 2. Illustration of the metabolomics work flow that combines advanced NMR spectroscopy techniques with multivariate statistics. Samples are collected in a uniform way to minimize variability and are analyzed by NMR profile to collect data on all metabolites potentially present in the sample. Pattern recognition approaches (**PRA**) include principal component analysis, partial least squares discriminant analysis, orthogonal projections to latent structures, heat map, support vector machines, and random forests method, and other modes aiming to highlight underlying trends and visualization tools such as contribution. Trend and box plots are used to further evaluate these. Receiver operating characteristic (**ROC**) curves are generally considered the method of choice for evaluating the performance of potential biomarkers. The markers are eventually placed in a metabolic pathway to provide insight on the biochemical phenomena. (**Samples**) Multiple replicate samples are obtained from cells, tissues, or biofluids (*e.g.,* urine, plasma, *etc.*) for each group (*e.g.,* healthy vs. disease). (**NMR profile**) A 1D ^{1}H NMR spectrum is collected for each metabolomics sample, which becomes the data set. A mass spectrum can be used instead or in addition to the NMR spectrum. (**PRA**) Illustrations of typical multivariate statistical analysis of the metabolomics NMR data set. Clockwise from upper-left, a scores plot from a PCA model indicating two distinct clusters or groups are present in the data set. A heatmap shows the clustering of metabolite changes (x-axis) relative to each group replicate (y-axis). The relative color of each bin corresponds to the metabolite concentration difference between replicates. The heatmap identifies which set of metabolites are uniquely changing between each group. The result of a random forest classification is summarized by plotting the out-of-bag error rate or misclassifications versus the number of trees. The results indicate that the healthy and disease groups can be separated with an error rate of <5% with a nominal number of trees. A back-scaled loadings plot, which is a pseudo 1D ^{1}H NMR spectrum, is generated from an OPLS model. The relative intensity and direction of each peak conveys the importance and correlation of the NMR peak to the observed group separation in the corresponding scores plot. (**ROC**) Illustrations of the validation and the further analysis of the key metabolites identified from the multivariate statistical analysis that define the group separation. On the left is a ROC curve, which plots sensitivity (true positive rate) versus 1-specificity (false positive rate). The area under the curve is a measure of the accuracy of the model to correctly predict group membership. On the right is the pathway topology analysis produced by MetPA (http://metpa.metabolomics.ca) from the list of metabolites identified by the multivariate statistical analysis. MetPA assists in identifying the set of important metabolic pathways associated with the phenotype. (**Pathways**) Illustration of metabolic networks or signaling pathways identified from the observed metabolomic changes between the groups. Reproduced with permission from Ref. [86].
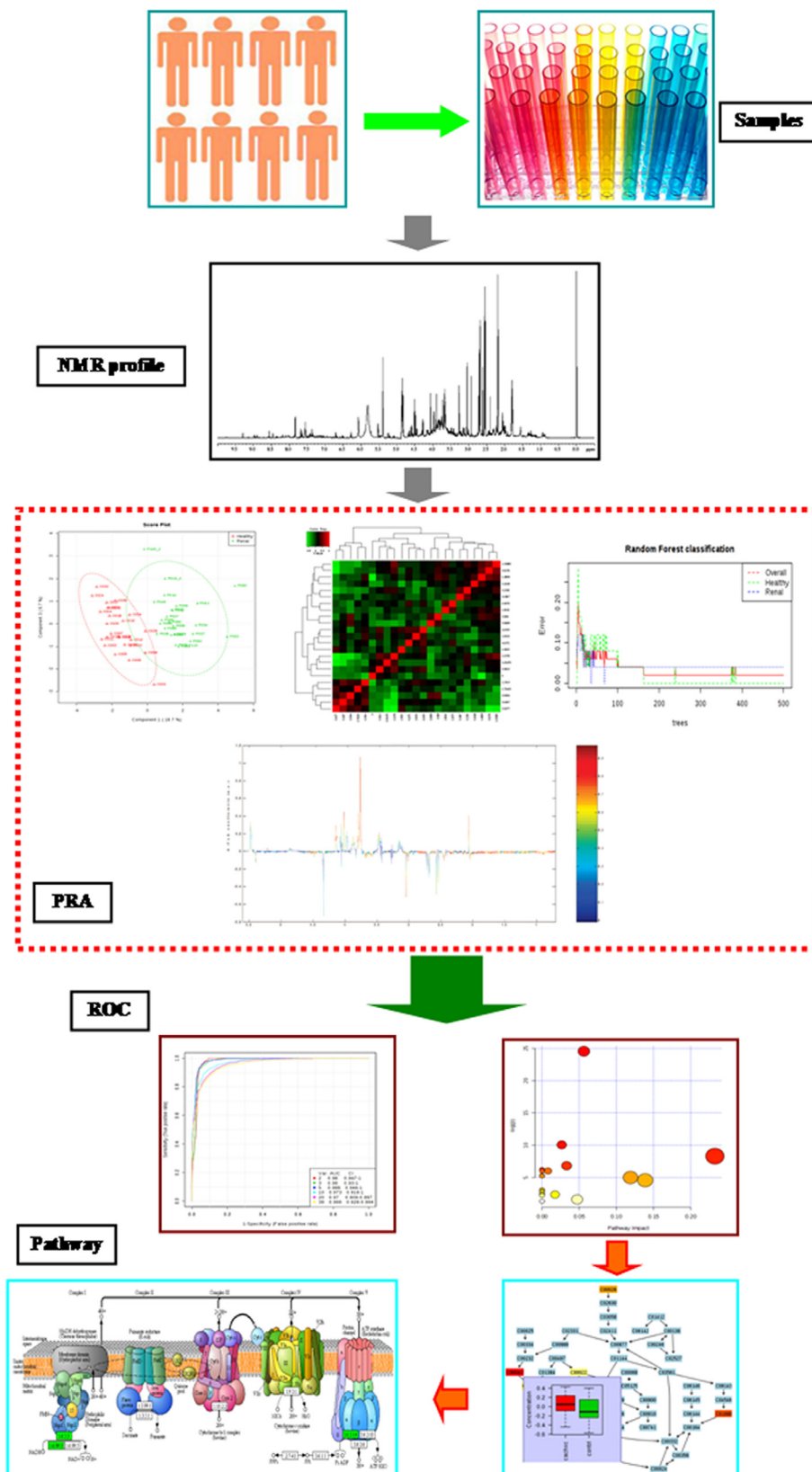
conditions, which depend on mode of polarization, pH and the ionization efficiency of the specific metabolite. Thus, the detectability of a metabolite may vary dramatically depending on sample conditions. For example, 0.1% formic acid or acetic acid is commonly added to MS samples to enhance protonation and increase sensitivity. Conversely, a deuterated solvent used in NMR could complicate the analysis of an MS spectrum because exchangeable hydrogens would still be observed, but ambiguous mass shifts may occur. The exchangeable hydrogens are likely not observed in the NMR spectrum, potentially eliminating the ability to detect a given metabolite. Moreover, minimal sample handling is desired to reduce experimental errors. Thus, these necessary differences in sample conditions present a practical challenge, namely how can a single metabolomics sample be analyzed by two distinct analytical platforms?

To address the need for distinct NMR and MS sample preparations, Beltran et al. examined 12 extraction protocols and evaluated various solvents and temperatures to identify conditions compatible for use by both LC-ESI-MS and NMR [101]. A single metabolite extract from a liver tissue was used for the NMR and MS metabolomics studies. The combination of solvent conditions used for metabolite extraction ranged in polarity and consisted of a 2 mL combination of either: (i) 1:1 methanol/H_2O, (ii) 1:1 acetonitrile/H_2O, (iii) 7:2:1 methanol/chloroform/H_2O, or (iv) 7:2:1 acetonitrile/chloroform/H_2O. The extractions were also conducted at three different temperatures: $-20\,°C$, $25\,°C$, or $60\,°C$. After each extraction, the metabolomics samples were dried and then dissolved in deuterated acetonitrile/H_2O (2:8) before acquiring a 1D ^1H NMR spectrum. A resulting PCA model comparing the range of solvent conditions revealed a large variance between the different extraction protocols. Interestingly, extraction temperature had a minor impact on group separation. Methanol/H_2O was the most efficient extraction method and the least influenced by temperature. In contrast, acetonitrile/chloroform/H_2O was the least efficient extraction method.

A very intriguing aspect of this study was the LC-MS analysis of metabolites in deuterated solvent. Pure metabolomics standards or liver metabolite extracts were prepared in either H_2O or D_2O (Fig. 3) [101]. In this manner, the impact of solvent exchange on NMR and MS spectral data was evaluated. D_2O did not affect measurements of metabolite concentrations. In addition, D_2O did not induce a mass shift or perturb the isotopic distribution. Metabolites with readily exchangeable hydrogens were also investigated and, surprisingly, no difference in the spectra due to D_2O was observed. The lack of deuterium exchange in metabolites was attributed to the mobile phase; since the deuterated metabolomics sample was injected into an un-deuterated mobile phase, deuterated metabolites were readily back-exchanged to a protonated species.

A similar study conducted by Marshall et al. identified a relatively straightforward and simple protocol to simultaneously investigate by both NMR and MS a metabolomics sample extracted from a cell lysate [57]. Human dopaminergic neuroblastoma cells were lysed by incubating with cold methanol ($-80\,°C$) for 15 min. The methanol was also used as the first metabolite extraction step. After removing the methanol supernatant, the cell debris were then washed with an 80%/20% methanol/ddH_2O water mixture, followed by a second wash with 100% ddH_2O. The supernatants from the three extractions were combined and then split into two portions: 1.8 mL for NMR and 200 μL for MS. The MS sample was then diluted tenfold with a 49.75:49.75:0.5 H_2O/methanol/formic acid mixture containing 20 μM reserpine as an internal mass reference. The NMR sample was dried using a combination of RotoSpeed vacuum and lyophilization, and then resuspended in a 50 mM deuterated phosphate buffer (pH 7.2) with TMSP-d_4 as a chemical shift reference. The separate NMR and MS metabolomics samples were then used to obtain a 1D ^1H NMR spectrum and a direct-injection ESI-MS mass spectrum.

3. Approaches for combining NMR and MS for metabolomics

3.1. Interfacing NMR and MS hardware

Natural product chemistry is a staple of drug discovery and the pharmaceutical industry, and has an obvious synergy with metabolomics since both fields are focused on the characterization of chemical entities extracted from a complex biological mixture [102,103]. NMR and MS spectral data have been routinely used in tandem to characterize natural products and their secondary metabolites isolated from crude complex samples [102]. Since metabolites and natural products are essentially the same, these well-established protocols are an invaluable resource for characterizing the vast number of unknown metabolites in the metabolome. Natural product research is commonly guided by and focused on identification of biologically active fractions from complex mixtures, which are then analyzed by a sequential series of purification steps and MS profiling until a biologically active compound is isolated for structural determination by NMR and MS methods [104,105]. In general, only a few nanograms or micrograms of a natural product are available for structural characterization. Therefore, sensitivity and efficiency are key concerns. In the case of NMR, the development of microcoils, microflow and cyroprobes have significantly improved the dynamic range and sensitivity of NMR and have thus greatly benefited the structural characterization of sample-limited natural products and metabolites [106–108]. Mass spectrometry has made similar advancements that have also benefited natural products chemistry and metabolomics by developing high-resolution mass spectrometers and tandem MS-MS methods to improve the accuracy and ease of structure elucidation [109–113].

The off-line combination of MS and NMR has clearly advanced natural products and metabolomics research, but an innovative on-line combination of the two instruments has the potential to significantly increase throughput, efficiency and sensitivity [59,60]. HPLC fractions are routinely collected and separately analyzed by NMR and MS, but an on-line analysis allows for the simultaneous detection of metabolite data that are then easily cross-correlated between the two platforms while also reducing, or potentially eliminating, sample handling errors. This is especially true in regard to the manual manipulation of severely sample-limited natural products or metabolites. Additionally, the duplicate isolation of redundant or unwanted compounds is a large issue in the natural products field. It is simply undesirable to expend valuable resources to re-discover known natural products. Dereplication is the commonly employed process of examining active fractions to recognize and eliminate previously identified natural products to avoid unnecessary compound isolation [114]. Dereplication routinely occurs by rapidly comparing experimental NMR and/or MS spectral data against spectral libraries of known natural products [60]. Employing an LC-NMR-MS instrument for dereplication would significantly reduce the time requirements, lower the amount of sample required and increase identification accuracy [60].

Previously, physically siting an instrument in close proximity to an NMR magnet was difficult, as it caused severe spectral distortions, required long sample transfer lines, and demanded significant laboratory space [59]. Fortunately, interfacing LC, NMR, and MS instrumentation has been greatly simplified by the advent of shielded magnets [115–117]. The stray fields have been significantly reduced for shielded magnets, and, in some cases, the 5 gauss line lies within the magnet's dewar. This permits a close
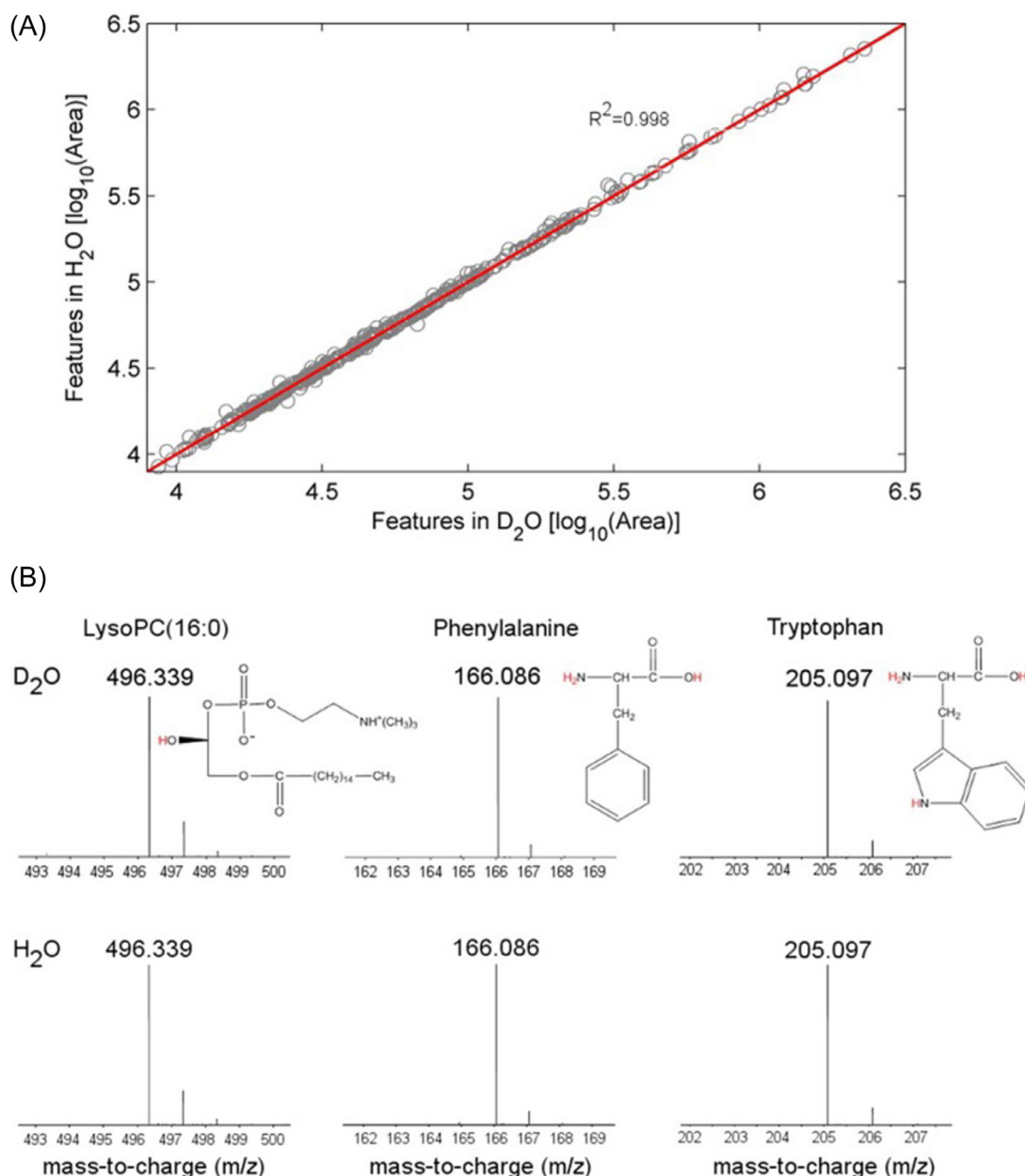
Fig. 3. (A) Scatter plot representing the area of each feature from the XCMS matrix of LC/MS data of liver samples reconstituted in H_2O and D_2O. A correlation coefficient (R_2) of 0.998 indicates a high linear regression, which demonstrates that differences between the number and abundance of features detected in liver extracts reconstituted in H_2O and D_2O are insignificant. (B) Mass spectra of phenylalanine, tryptophan, and LysoPC (16:0) reconstituted in D_2O (top) and H_2O (bottom). Labile hydrogens are marked in red. Mass spectra show that the isotopic distributions of the compounds are not altered by D_2O, indicating either slow H/D exchange of acidic protons in solution or fast back-exchange of labile deuterons in aqueous LC/MS buffers due to a total solvent accessibility of small molecule structures. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) Reproduced with permission from Ref. [101].

arrangement of instruments without incurring any magnetic interference or disturbing field homogeneity. The availability of bench-top mass spectrometers has also aided in the coupling of multiple platforms. An MS instrument with a small footprint and superior resolution makes it possible to site an LC-NMR-MS spectrometer in a smaller laboratory space then previously possible. Moreover, the proximity of the instruments simplifies sample transfer between them. For example, an NMR and mass spectrometer can be interfaced to LC by using a post-column splitter or the Bruker NMR-Mass Spectrometry Interface (BNMI). The post-column splitter simultaneously directs LC flow to both the NMR and MS spectrometers in a 9–1 ratio, respectively. BNMI is a valve-switching interface that is a computer-controlled splitter and double dilutor, which allows for proton–deuterium exchange to occur for NMR,

and the optimal selection of solvent polarity for MS. Additionally, BNMI has a loop storage mode, which allows for a portion of the LC eluent to be temporarily stored in a sample loop while waiting for the NMR instrument to be available. A storage mode allows for NMR acquisition times that are significantly longer than the LC experiment to maximize NMR signal-to-noise.

Recently, Lin et al. described an LC-NMR-MS approach, which significantly reduced sample size, increased concentration sensitivity by 10-fold, and enhanced mass sensitivity by 1000-fold [59,98]. The method significantly outperformed traditional offline LC-NMR-MS by the addition of a nanoSplitter for nanospray LC-MS and a microcoil flow probe for NMR analysis. In addition, the analysis of the LC flow normally occurs with a 9:1 split between NMR and MS, respectively. But, instead, the improved performance

required an LC flow that was split 98% to NMR and 2% to MS [59]. The LC-NMR-MS system exhibited a 93% sample recovery and yielded a limit of detection as low as 50-ng (RSD 1.17%) for NMR. Similarly, four natural products (ambiguine A, I, E, and hapalindole H) from a cyanobacterial extract were readily detected by LC-MS at a concentration of only 30 μg/mL. Thus, interfacing LC with parallel NMR and MS analysis may greatly benefit metabolomics by improving throughput, by increasing sensitivity and coverage, and by reducing the amount of sample required for a study.

3.2. Chemical modification

The diverse and complementary set of information obtained from NMR and mass spectral data is the major reason for combining the two methods for metabolomics. However, combining two distinct data sets also increases the complexity of the analysis. The NMR and MS results obtained from a single heterogeneous sample are not easily correlated [30]. For example, it is not a trivial task to definitively assign an NMR chemical shift and a MS m/z value to a specific metabolite. Simply, there is no information in the NMR or MS spectrum that indicates that the chemical shift and m/z value are from the same metabolite. The lack of a correlation may also appear to produce contradictory results. For example, the presence of a metabolite may only be supported by one method or the two methods may appear to predict vastly different concentrations. As a result, the complexity of analyzing NMR and MS spectra may hinder an interpretation or yield erroneous results. Consequently, a novel twist on the well-established method of chemical derivation has been employed to overcome this lack of a correlation between NMR and MS spectral data [30]. A chemical

agent is introduced into a complex mixture that reacts with a specific chemical moiety or functional group within the metabolite. The resulting chemical modification is then visible by both NMR and MS.

Chemical modification has frequently been used for LC-MS and GC–MS [118–121]. Traditionally, compound derivatization has been used to separate stereoisomers or to improve ionization efficiency in LC-MS. Chemical modification for GC–MS is generally used to promote the volatility of compounds [122]. Fariba et al. exploited the benefits of chemical modification by using ^{15}N-cholamine, a so-called "smart tag", to specifically label carboxyl-containing metabolites [30,123]. A permanent charge and an ^{15}N isotope label are incorporated into the metabolites through a condensation reaction [30]. Each chemically labeled metabolite would then have a unique set of ^{1}H and ^{15}N chemical shifts that can be easily detected with a 2D ^{1}H-^{15}N HSQC NMR spectrum. Only metabolites labeled with ^{15}N-cholamine are observed in the NMR spectrum because of the low natural abundance of ^{15}N (0.37%). Similarly, the incorporation of a permanent charge into the metabolite significantly improved its ionization efficiency. Sensitivity enhancements exceeding three orders of magnitude were observed for labeled metabolites, which could be easily detected in the positive mode compared to unlabeled metabolites in the negative mode. Thus, the increased MS sensitivity and the unique ^{1}H-^{15}N chemical shifts establish a correlation between the NMR and MS data for a labeled metabolite. The smart-tag approach was successfully demonstrated with a standard mixture of 48 metabolites that each contain a carboxyl-group and are prevalent in human biofluids (Fig. 4) [30]. The experiment was successfully repeated with human serum and urine samples. The chemical shift
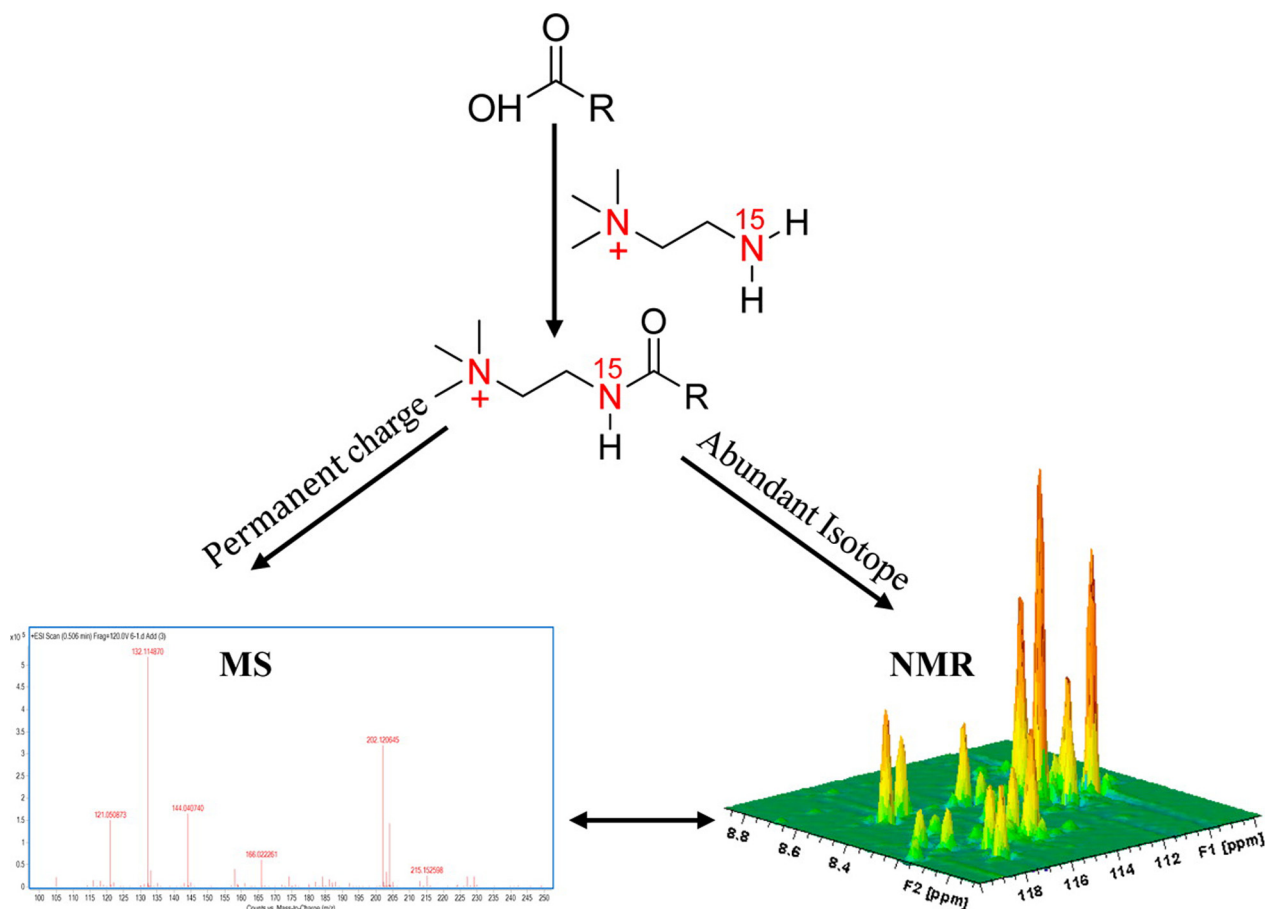

Fig. 4. Schematic figure illustrating the "smart isotope tag" approach used to detect the same metabolites using NMR and MS with high sensitivity. Tagging carboxyl-containing metabolites with ^{15}N-cholamine enables their enhanced detection by both NMR and MS. Reproduced with permission from Ref. [30].

assignments obtained from the standard metabolite mixture was used to quickly and accurately assign the 48 metabolites in the biofluids. Thus, ^{15}N-cholamine was shown to be useful for correlating data across multiple analytical platforms through chemical derivatization, and as a valuable approach to aid in the metabolite profiling of complex biological mixtures.

3.3. Stable isotope tracers

The ability to simultaneously measure hundreds of chemical species from a complex biological sample is the primary reason NMR and MS are routinely used in metabolomics, but the overwhelming amount of data can also confound the analysis. Consequently, the majority of metabolomics studies are end-point or single time-point measurements despite the obvious advantage of flux analysis. Fortunately, MS and NMR can de-convolute these complex data sets based on their unique ability to identify and distinguish between different isotopes. Thus, NMR and MS metabolomics analysis can be greatly simplified by using stable isotope tracers. Stable Isotope-Resolved Metabolomics (SIRM) has been used to monitor metabolite flux, to reveal novel metabolic networks, and has recently been shown to correlate data across MS and NMR platforms [124–130]. In this manner, SIRM may overcome common limitations encountered with steady-state metabolite profiling. SIRM also enables the combined use of NMR and MS for metabolomics, but, more importantly, the SIRM approach greatly benefits from combining both NMR and MS [126–129,131].

SIRM uses stable isotope-enriched nutrients containing ^{13}C-carbons or ^{15}N-nitrogens, which are provided to an organism as the primary source of carbon or nitrogen. The ^{13}C-carbons or ^{15}N-nitrogens are distributed throughout the organism's metabolome based on active metabolic processes. The resulting ^{13}C- or

^{15}N-labeled metabolites are readily detected by NMR and MS. In this manner, SIRM reveals the unique flow of ^{13}C-carbons or ^{15}N-nitrogens through metabolic pathways in order to highlight systematic changes due to environmental stress, genetic mutations, a disease state, a drug treatment, or numerous other factors. Importantly, SIRM allows for monitoring the flow of position-specific carbons or nitrogens between metabolites (i.e., isotopomer or isotopologue probing). In essence, it is possible to decipher the chemical source of each carbon or nitrogen in a given metabolite (Fig. 5) [132]. Moreover, the analysis can be fine-tuned by repeating the experiments with different labeled metabolites (e.g., glucose, pyruvate, etc.) or different labeling within a metabolite (e.g., [1-^{13}C] glucose, [2-^{13}C] glucose, etc.). Identification of position-specific carbon or nitrogen labeling is relatively straightforward by NMR since, in general, each carbon or nitrogen has a unique chemical shift. For example, consider the distinct carbon chemical shifts observed for L-alanine: 178.5 ppm (C′), 53.4 ppm (Cα) and 19.0 ppm (Cβ). Conversely, it is not possible to identify position-specific carbon or nitrogen labeling by MS except in the case of a compound containing a single carbon or nitrogen, or by employing MS^n fragmentation analysis. Instead, the number of isotopically labeled carbons or nitrogens is easily determined by MS from an observed mass shift. In addition, the abundance of each isotopomer is readily obtainable from an MS spectrum by comparing the relative intensity of each molecular-ion peak. Conversely, obtaining the number and abundance of each isotopomer and isotopologue is not easily obtained from an NMR spectrum.

Lane et al. used direct infusion FT-ICR-MS combined with 1D and 2D NMR methods to identify isotopomers of glycerophospholipids (GPL) derived from [U-^{13}C]-glucose in breast cancer MCF7-LCC2 cell extracts [124]. An algorithm was also presented that accounts for the contribution of natural abundance ^{13}C following the incorporation of ^{13}C-carbon derived from glucose [133]:
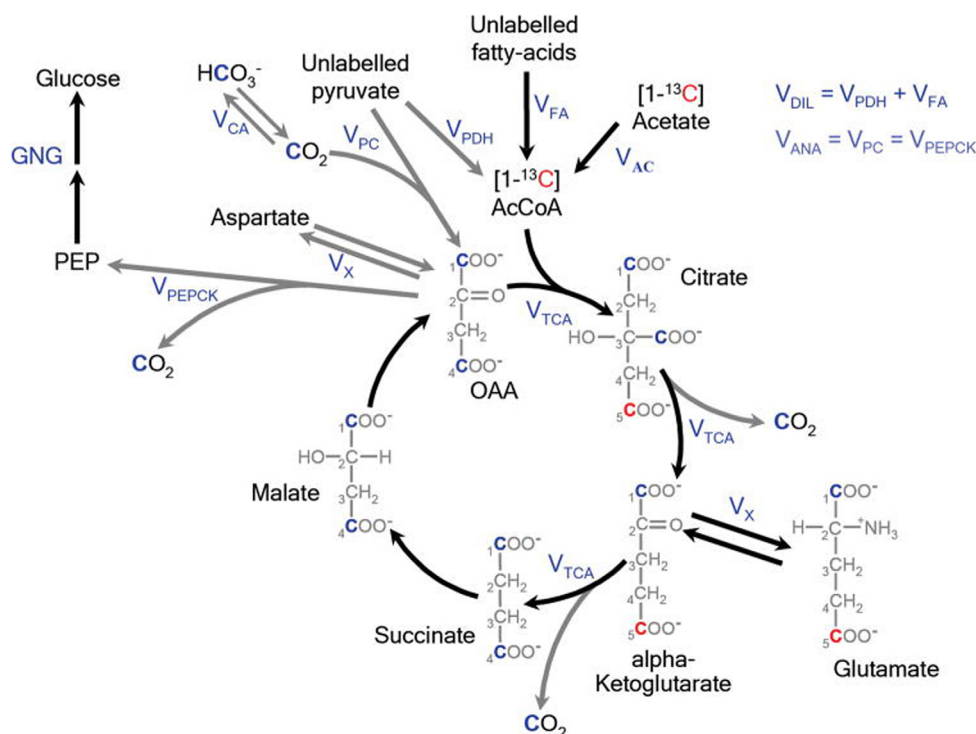

Fig. 5. Metabolic model of liver acetate oxidative metabolism used to estimate hepatic TCA cycle flux (V_{TCA}) and anaplerosis (V_{ANA}). Carbon positional enrichment denoted in red occurs during the initial incorporation of label from 1-^{13}C-acetate to glutamate on the first pass through the TCA cycle. Positional enrichment denoted in blue occurs during the 2nd pass through the TCA cycle, with label originating from internal scrambling at succinate or from bicarbonate (HCO$_3^-$)/$^{13}CO_2$ via anaplerosis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) Reproduced with permission from Ref. [145].
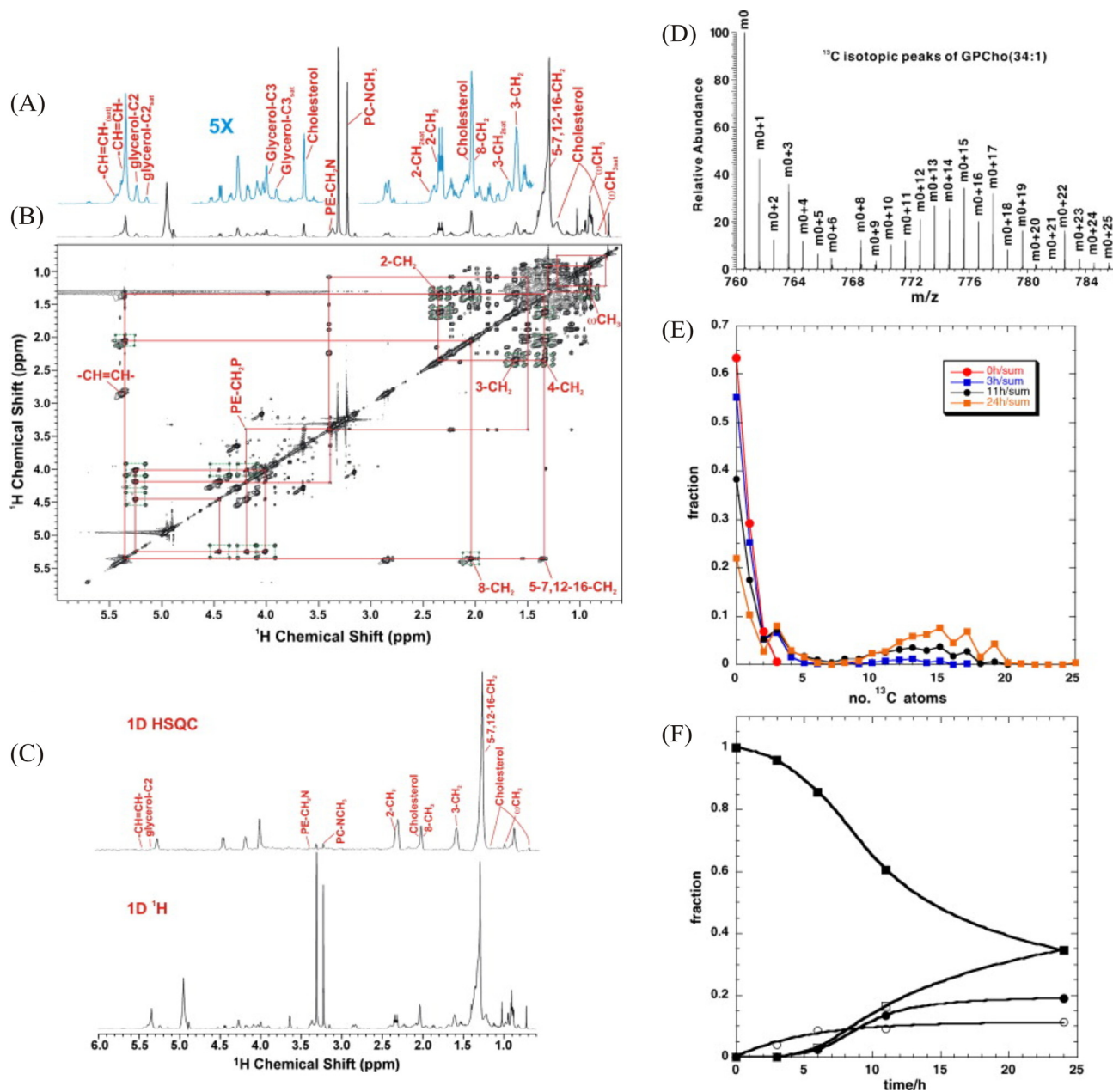
Fig. 6. High resolution NMR spectra of a methanolic extract of LCC2 cells. Glycerophospholipids were extracted from LCC2 cells grown in the presence of 10 mM [U-^{13}C]-glucose for 24 h. (**A**) 1D ^{1}H NMR spectrum and (**B**) TOCSY spectrum. The TOCSY spectrum was recorded at 18.8 T 293 K with 50 ms mixing time at a B$_1$ field strength of 9 kHz. The data were processed with one linear prediction and zerofilling in t$_1$ and apodized using an unshifted Gaussian function in both dimensions. (**C**) 1D ^{1}H NMR spectra, top: 1D ^{13}C-edited ^{1}H (HSQC) spectrum, bottom: high resolution ^{1}H NMR spectrum. High resolution FT-ICR mass spectrum of a methanolic extract of LCC2 cells. (**D**) FT-ICR-MS profile spectrum of an LCC2 methanol extract after 24 h labeling with [U-^{13}C]-glucose. A close up of the m/z region from 760 to 782 is shown. The accurate masses (better than 1 ppm) at high resolution (>100,000 at measured mass) enable assignment of the GPLs and their isotopologues. Masses were externally calibrated, and secondarily calibrated with respect to internal standard reserpine; intensities have been arbitrarily scaled to 100 units for m$_0$ at m/z = 760.5860. (**E**) Mass distribution of PC 34:1 normalized to the total intensity as a function of time. The distribution at 0 h is indistinguishable from the expected natural abundance intensity. Line graphs are used here for clarity only; no values are implied between data points. (**F**) Time courses of selected mass peaks. (■) m$_0$, (□) m$_0$ + 3, (●) Σ(m$_0$ + 2n); (○) Σ(m$_0$ + 3 + 2n). The m$_0$ + 3 intensities were fitted to a $(1 - \exp(-kt))$ with a = 0.11 ± 0.008 and k = 0.19 ± 0.04 h^{-1}. Reproduced with permission from Ref. [124].

$$I_{M+i} = \frac{I_{M+i:NA} - \sum_{x=0}^{x<i} I_{M+x} * B_C(x,i)}{1 - B_c sum(i)};$$

$$B_c sum(i) = \sum_{k=n+1}^{C_{Max}} B_c(n,k);$$

$$B_C(n,k) = \binom{C_{Max} - n}{k - n}(1 - NA_{13_C})^{C_{Max}-k} NA_{13_C}^{k-n}$$

where $I_{M+i:NA}$ is the expected intensity of the ith isotopologue peak, NA_{13_C} is the ^{13}C natural abundance (1.1%), C_{Max} is the total number

of carbons in the molecule, k is the total number of ^{13}C carbons, n is the number of ^{13}C carbons incorporated from a labeling source, k-n is the number of natural abundant ^{13}C carbons. Analysis of the 1D ^{1}H NMR spectra identified phosphatidylcholines with approximately two double bonds as the major GPL present in the MCF7-LCC2 cell extracts (Fig. 6A–C). Based on the relative intensity of the assigned GPL NMR resonances, it was determined that the choline head groups were not ^{13}C labeled, but the fatty acyl chains and the glycerol moieties were predominately derived from [U-^{13}C]-glucose. Specifically, the glycerol moiety was determined to be

44 ± 1.2% ^{13}C, and the C2, C3 and C4 positions of the fatty acids were determined to have an average ^{13}C incorporation of 46 ± 4%. These results indicated that nearly 50% of GPL was newly synthesized in 24 h. The FT-ICR-MS data (Fig. 6D–F) complemented the NMR results by identifying the individual GPL species [*i.e.,* PC (34:1)] and the abundance of the corresponding ^{13}C isotopologues as a function of time. Specifically, the m_0 isotopologue decreased and the $m_0 + 3$, $m_0 + 2n + m_0 + 3 + 2n$ isotopologues sequentially increased as ^{13}C-carbon from glucose was incorporated into PC (34:1). The odd number isotopologues were derived from ^{13}C glycerol and unlabeled fatty acids, while conversely; even number isotopologues resulted from unlabeled glycerol and ^{13}C fatty acids. Interestingly, the abundance of the n = 10–20 fatty acid chains increased with time, but only a minimal mass shift was observed. A ^{12}C acetate pool persisted for twenty-four hours from fatty acid turnover and internal triglyceride stores. In fact, only 30 to 50% of the 34 carbons were incorporated into PC (34:1) from ^{13}C acetyl CoA. Consistent with the NMR results, the mass spectral data indicates that 46% of glycerol and 44% of the fatty acyl moieties are ^{13}C-labeled. Moreover, based on these measurements, isotopologue distributions of GPLs over a twenty-four period were accurately measured and simulated. Thus, the authors clearly demonstrate the value of combining NMR and MS to monitor the synthesis of various GPLs from pools of metabolite. The combination of MS and NMR yielded positional isotope labeling information, ^{13}C isotopologue distributions, and enabled the accurate and efficient identification of GPL species. SIRM using both NMR and MS greatly improves the analysis of metabolomic flux (*i.e.,* metabolite synthesis and turnover).

3.4. Combined cheminformatics methods

The effective handling and analysis of large amounts of information (*i.e.,* "Big data") presents unique challenges [134–137]. The field of metabolomics faces similar concerns, which make it difficult to handle the large amounts of information with traditional methods such as database management, basic statistical methods, or simple manual analysis [138]. Combined cheminformatic approaches present a valuable alternative to accelerate the accurate processing of "omics" data sets. Since a considerable amount of data may be generated by combining NMR and MS, combined cheminformatics approaches are increasingly being employed during metabolomics studies. Recently, the SUMMIT and NMR/MS translator methods were developed for the rapid and accurate identification of metabolites [28,29]. SUMMIT MS/NMR and the NMR/MS translator combine NMR and MS to elucidate the structures of unknown metabolites from complex mixtures. NMR/MS translator combines COLMAR [139] database search queries and experimental NMR and MS spectral data to accelerate accurate metabolite identification [28]. NMR/MS translator uses 1D ^{1}H NMR, or 2D ^{1}H–^{13}C HSQC and 2D ^{1}H–^{13}C HSQC TOCSY chemical shifts to perform a COLMAR database search and return a list of possible metabolite candidates. The query candidate list is then used to produce a simulated mass spectrum for each possible metabolite, which includes possible adducts, fragments, and isotope distributions. These simulated mass spectra are then compared against an experimental metabolomics mass spectrum to make metabolite assignments. In effect, potential metabolites identified by NMR are confirmed by MS. As a proof of concept, the NMR/MS translator was used to successfully analyze a model mixture of 26 metabolites with 2D ^{1}H-^{13}C HSQC NMR spectra and DI-ESI-MS positive and negative mode spectra. The NMR/MS translator was further validated using a set of human urine samples from healthy volunteers. A total of 98 urine metabolites were identified by the NMR/MS translator, which included 8 metabolites that were not previously observed in a comprehensive study of

human urine. Importantly, only 48 of these metabolites were correctly identified using MS data alone, including MS/MS fragmentation patterns. The NMR/MS translator approach automates the metabolite assignment and avoids labor-intensive manual analysis, which enhances coverage, improves consistency, and increases throughput.

SUMMIT MS/NMR is an alternative high throughput approach that combines MS with NMR data to identify unknown metabolites in a complex biological sample (Fig. 7) [29]. The SUMMIT MS/NMR approach relies on the acquisition of a high-resolution mass spectrum, in which each m/z peak is converted into a molecular formula. The list of molecular formulas is then used to generate a set of all feasible structures (*e.g.,* a structural manifold) with the ChemSpider database [140], which are used to predict an NMR spectrum using MestReNova 9.0.1 (Mestrelab Research, Santiago de Compostela, Spain). The COLMAR algorithm [139] is then used to compare the experimental NMR data against the database of predicted NMR spectra. Importantly, the experimental 2D ^{1}H–^{13}C HSQC NMR spectrum is deconvoluted into subspectra corresponding to the individual components of the mixture using connectivity information derived from 2D ^{1}H–^{1}H TOCSY, 2D ^{1}H–^{13}C HSQC-TOCSY, and 2D ^{1}H–^{13}C HMBC spectra. The potential metabolites are rank-ordered based on the relative agreement between the experimental NMR subspectrum and the simulated NMR spectra.

SUMMIT MS/NMR was validated using a DI-ESI-MS spectrum of a model mixture containing 10 metabolites [29]. The 50 largest m/z peaks were selected, which yielded 22 molecular formulas, 362 potential structures, and 4772 predicted 2D ^{1}H–^{13}C HSQC NMR spectra. SUMMIT MS/NMR ranked 6 of the 10 metabolites as the top hit and three other metabolites were identified as the second-best hit. The remaining metabolite did not ionize. For the three metabolites identified as second best, the top hits were structurally very similar (*i.e.,* allo-isoleucine instead of leucine). The approach was repeated using an MS spectrum of a polar extract of an *E. coli* lysate. The 500 largest m/z peaks were selected corresponding to 56 molecular formulas and 13,872 structures and ^{1}H-^{13}C HSQC spectra. A total of 21 metabolites were accurately and rapidly identify by SUMMIT MS/NMR, and then confirmed using a set of 2D ^{1}H-^{13}C NMR experiments. Thus, both SUMMIT MS/NMR and NMR/MS translator clearly illustrate the inherent value of combining NMR and MS to enhance metabolomics.

3.5. Multivariate statistical methods

A well-known problem with metabolomics data is the presence of confounding factors that may complicate the identification of group membership. For example, the analysis of urine or serum to identify biomarkers may be masked by metabolites associated with age, diet, ethnicity, gender, or race, among other factors. Multivariate statistical methods are able to cope with these multiparametric data sets and extract group membership [141]. For a detailed and comprehensive review of multivariate statistics and its application for metabolomics see Worley et al. [12]. Chemometric techniques can be divided into supervised or unsupervised methods. In metabolomics, unsupervised methods are commonly used to identify global trends or group membership. Alternatively, supervised methods are highly valuable for identifying the spectral features (or metabolites) that primarily contribute to the differentiation between groups. PCA, PLS, and OPLS are the chemometric methods commonly used in metabolomics based on a *single* analytical source. The limited availability of chemometric methods applicable to multiple analytical sources is one reason NMR and MS have not been commonly combined for metabolomics.

Nevertheless, multivariate statistical techniques have been previously applied for the combined analysis of MS and NMR data sets.
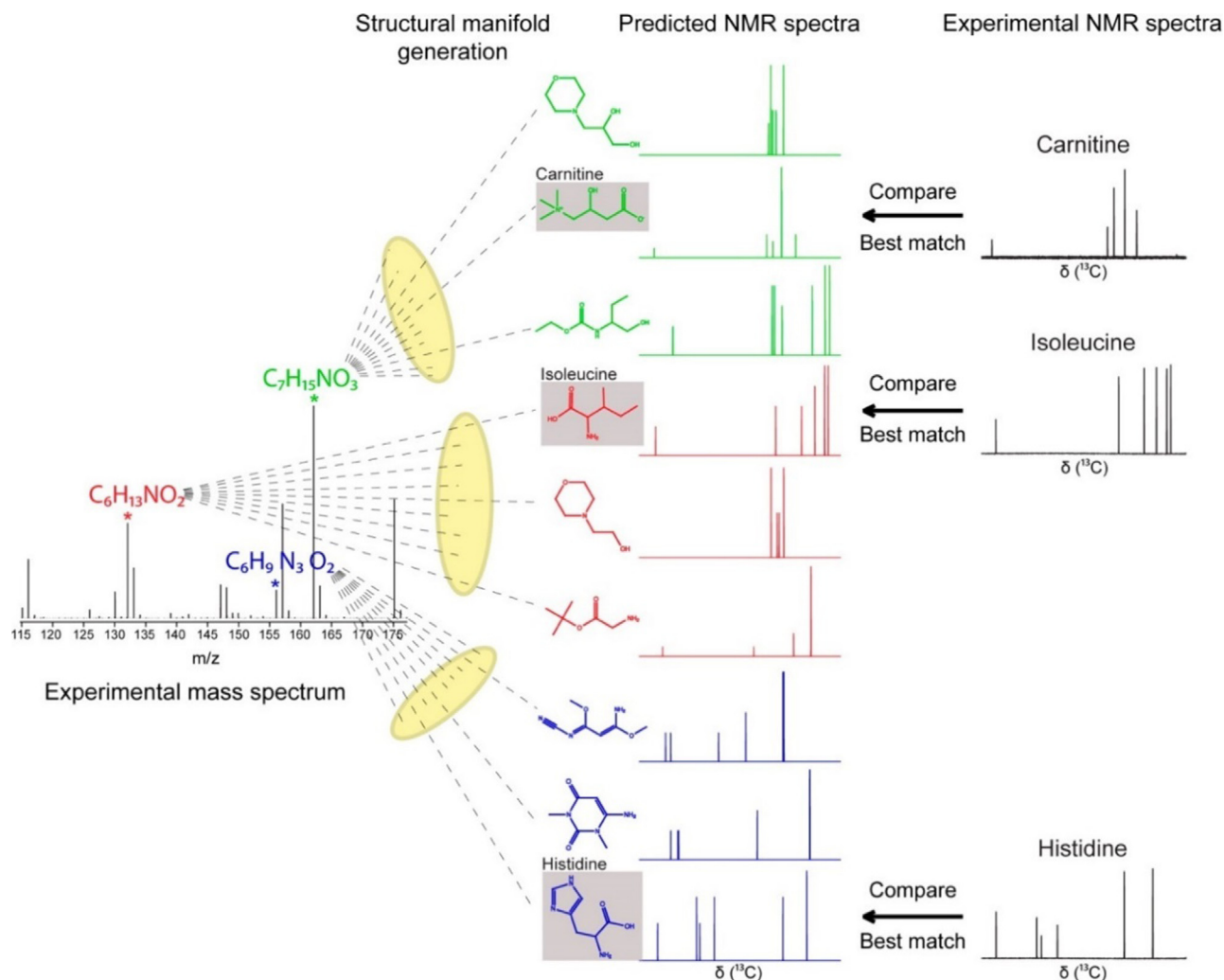
Fig. 7. Schematic representation of the SUMMIT MS/NMR strategy for the identification of metabolites in complex metabolomic mixtures by the combined use of mass spectrometry and 1D ^{1}H NMR spectroscopy. High-resolution MS yields the unique molecular formulas of the metabolites present in the mixture (left). For each molecular formula, all possible structures are generated, representing the total structural manifold depicted as the sum of the three local manifolds (green, red, blue; middle), each belonging to a different mass. Next, NMR chemical shifts are predicted for all manifold structures. Comparison of the predicted with the experimental NMR chemical shifts (right) allows identification of the structures that are present in the mixture, requiring neither an NMR nor an MS metabolomics database [28,29,58]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) Reproduced with permission from Ref. [29].

Chen et al. generated individual PCA models for NMR and MS data sets and then combined the scores from each analysis into a three dimensional (3D) scores plot. The combined scores yielded a greater between-class separation than the original NMR or MS scores alone. Unfortunately, such an analysis ignores the highly informative correlations that exist between the two data sets. Gu et al. replaced the binary class designation of an MS data set with the first principal component (PC1) from a PCA model generated from NMR data to produce a subsequent OPLS-DA model for the MS data [70]. Again, a greater class separation was observed when the MS OPLS-DA model was generated with the NMR PC1 compared to the binary classification [57]. Nevertheless, such an analysis carries no statistical guarantee of success for any data set.

Recently, it was shown that a chemometrics model generated by integrating NMR and MS metabolomics data provided better group separation and a greater level of model interpretability than with NMR or MS data sets alone [57]. Marshall et al. combined 1D ^{1}H NMR and DI-ESI-MS. Multiblock methods are similar to traditional PLS and PCA, but provide a means for analyzing data from multiple analytical sources [142–144]. The spectral observations from each analytical method are placed into separate "blocks,"

which allows for the generation and simultaneous usage of within-block and between-block data correlations. Since the blocks share common trends, a model based on the between-block correlations will provide a better agreement with the biological groups. In effect, better discrimination between groups is expected by combining NMR and MS data than would be achieved from only the individual data sets. DI-ESI-MS and 1D ^{1}H NMR spectra were collected on cell lysates obtained from human dopaminergic neuroblastoma cells (SK-N-SH) treated with different neurotoxins: rotenone, 6-hydroxydopamine (6-OHDA), 1-methyl-4-phenylpyridinium (MPP$^+$), or paraquat. The PCA model produced from the 1D ^{1}H NMR data set yielded only two groups corresponding to the untreated controls and cells treated with the different neurotoxins. In effect, NMR detected no difference in the metabolome of neuronal cells after treatment with the different neurotoxins. The PCA model generated from the MS data set produced a modest separation between four groups. The untreated controls, MPP$^+$, and paraquat treatment each formed a separate group. Both the rotenone and 6-OHDA cell treatment were clustered together and formed the fourth group (Fig. 8). The MB-PCA and MB-PLS models generated from both the DI-ESI-MS and 1D ^{1}H NMR data
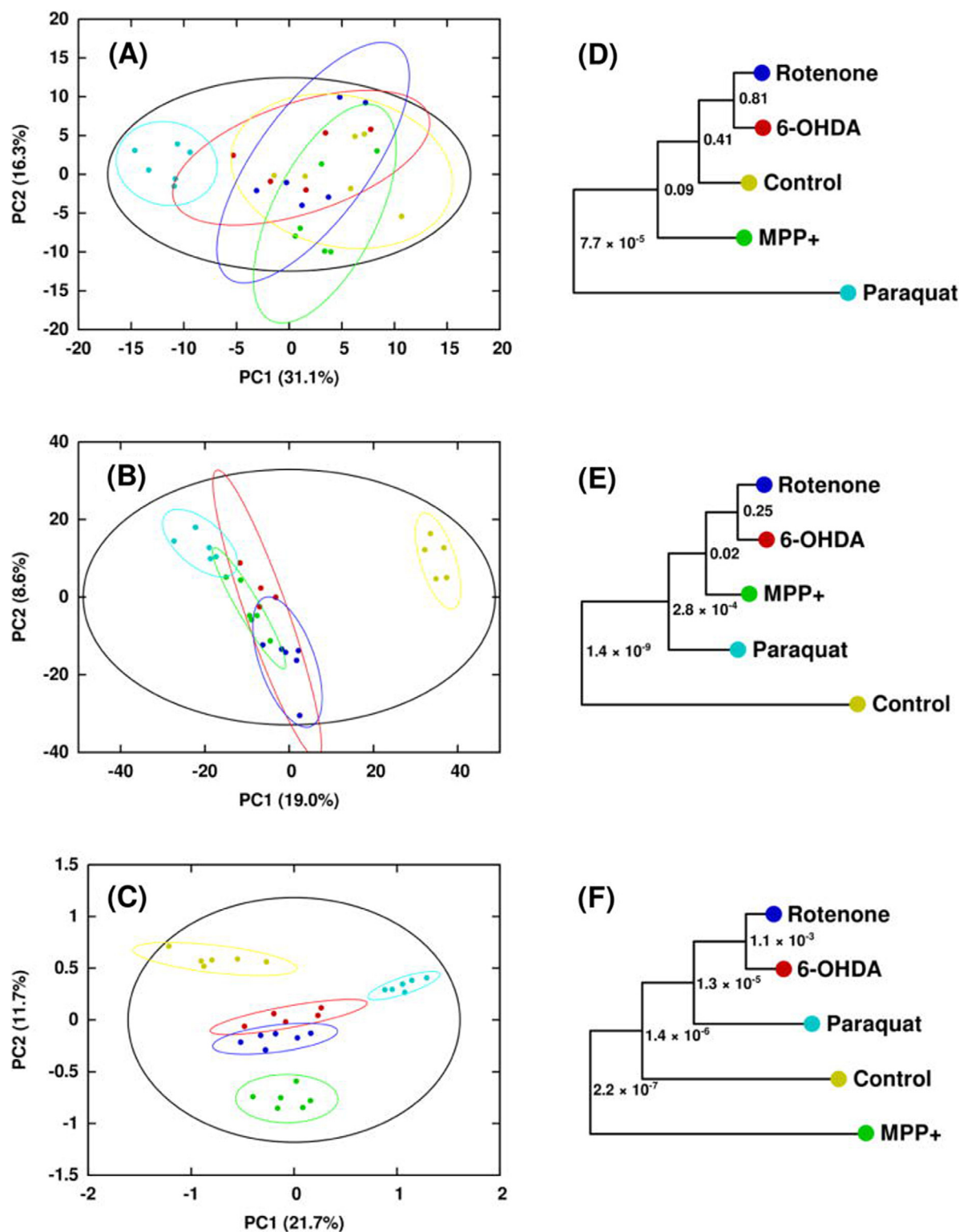
Fig. 8. Scores generated from (A) PCA of ^{1}H NMR, (B) PCA of DI-ESI-MS, and (C) MB-PCA of ^{1}H NMR and DI-ESI-MS. Separations between classes are greatly increased upon combination of the two data sets via MB-PCA. Symbols designate the following classes: Control (●), Rotenone (●), 6-OHDA (●), MPP$^+$ (●), and Paraquat (●). Corresponding dendrograms are shown in (D–F). The statistical significance of each node in the dendrogram is indicated by a *p* value. Reproduced with permission from Ref. [57].

sets yielded five distinct groups corresponding to each neurotoxin treatment and the untreated controls. This clearly demonstrated that each neurotoxin induced dopaminergic neuronal cell death through a distinct molecular mechanism. A detailed analysis of the metabolic impact of paraquat revealed that paraquat "hijacks"

the pentose phosphate pathway (PPP) to increase NADPH-reducing equivalents and stimulate paraquat redox cycling, oxidative stress, and cell death [57,71]. Thus, a successful outcome for a metabolomics study was critically dependent on combining NMR and MS data.

4. Conclusion

Metabolomics is an invaluable tool of systems biology and has made significant contributions to several diverse fields, including drug discovery, disease diagnosis, nutrition, environmental studies, and personalized medicine. To date, the majority of metabolomics data sets have been acquired using either MS or NMR separately. However, it is well known that combining MS and NMR data greatly improves the coverage of the metabolome and enhances the accuracy of metabolite identification. Consequently, combining NMR and MS techniques for metabolomics is a growing trend that will greatly benefit the quality and accuracy of metabolomics data. Herein we have reviewed several methodologies for integrating NMR and MS for the analysis of metabolomics samples. As demonstrated throughout this review, combining NMR and MS greatly enhances and improves the outcomes of metabolomics studies.

Acknowledgements

References

[1] S. Wachtel-Galor, I.F.F. Benzie, Herbal Medicine: An Introduction to Its History, Usage, Regulation, Current Trends, and Research Needs, 2011.
[2] S.G. Oliver, M.K. Winson, D.B. Kell, F. Baganz, Systematic functional analysis of the yeast genome, Trends Biotechnol. 16 (1998) 373–378.
[3] J.K. Nicholson, J.C. Lindon, E. Holmes, 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data, Xenobiotica 29 (1999) 1181–1189.
[4] O. Fiehn, Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks, Comp. Funct. Genom. 2 (2001) 155–168.
[5] W.B. Dunn, N.J. Bailey, H.E. Johnson, Measuring the metabolome: current analytical technologies, Analyst 130 (2005) 606–625.
[6] L.R. Bijland, M.K. Bomers, Y.M. Smulders, Smelling the diagnosis A review on the use of scent in diagnosing disease, J. Med. 71 (2013).
[7] J.L. Wolfender, G. Marti, A. Thomas, S. Bertrand, Current approaches and challenges for the metabolite profiling of complex natural extracts, J. Chromatogr. A 1382 (2015) 136–164.
[8] K. Bingol, R. Brüschweiler, Multidimensional approaches to NMR-based metabolomics, Anal. Chem. 86 (2014) 47–57.
[9] V.A. Likic, M.J. McConville, T. Lithgow, A. Bacic, Systems biology: the next frontier for bioinformatics, Adv. Bioinform. (2010) 268925.
[10] D.S. Wishart, Current progress in computational metabolomics, Brief. Bioinf. 8 (2007) 279–293.
[11] K. Dettmer, P.A. Aronov, B.D. Hammock, Mass spectrometry-based metabolomics, Mass Spectrom. Rev. 26 (2007) 51–78.
[12] B. Worley, R. Powers, Multivariate Analysis in Metabolomics, Curr. Metabolom. 1 (2013) 92–107.
[13] T. Gebregiworgis, R. Powers, Application of NMR metabolomics to search for human disease biomarkers, Comb. Chem. High Throughput Screen 15 (2012) 595–610.
[14] O. Fiehn, Metabolomics - the link between genotypes and phenotypes, Plant Mol. Biol. 48 (2002) 155–171.
[15] G. Astarita, J. Langridge, An emerging role for metabolomics in nutrition science, J. Nutrigenet. Nutrigenom. 6 (2013) 181–200.
[16] D.P. Jones, Sequencing the exposome: a call to action, Toxicol. Rep. 3 (2016) 29–45.
[17] S.J. Kim, S.H. Kim, J.H. Kim, H.J. Yoo, S. Hwang, Understanding metabolomics in biomedical research, Endocrinol. Metab. (Seoul) 31 (2016) 7–16.
[18] R.D. Beger, W. Dunn, M.A. Schmidt, S.S. Gross, J.A. Kirwan, M. Cascante, L. Brennan, D.S. Wishart, M. Oresic, T. Hankemeier, D.I. Broadhurst, A.N. Lane, K. Suhre, G. Kastenmuller, S.J. Sumner, I. Thiele, O. Fiehn, R. Kaddurah-Daouk,

Metabolomics enables precision medicine: "A White Paper, Community Perspective", Metabolomics 12 (2016) 1–15.
[19] G.A.N. Gowda, D. Raftery, Can NMR solve some significant challenges in metabolomics?, J Magn. Reson. 260 (2015) 144–160.
[20] G.A.N. Gowda, D. Raftery, Biomarker discovery and translation in metabolomics, Curr. Metabolom. 1 (2013) 227–240.
[21] J.C. Lindon, J.K. Nicholson, The emergent role of metabolic phenotyping in dynamic patient stratification, Expert Opin. Drug Metab. Toxicol. 10 (2014) 915–919.
[22] E.P. Rhee, R.E. Gerszten, Metabolomics and cardiovascular biomarker discovery, Clin. Chem. 58 (2012) 139–147.
[23] J.L. Markley, R. Bruschweiler, A.S. Edison, H.R. Eghbalnia, R. Powers, D. Raftery, D.S. Wishart, The future of NMR-based metabolomics, Curr. Opin. Biotechnol. 43 (2017) 34–40.
[24] T. Hirschfeld, The hy-phen-ated methods, Anal. Chem. 52 (1980) 297A–312A.
[25] A. Zhang, H. Sun, P. Wang, Y. Han, X. Wang, Modern analytical techniques in metabolomics analysis, Analyst (Cambridge, UK) 137 (2012) 293–300.
[26] G.A. Barding, S. Beni, T. Fukao, J. Bailey-Serres, C.K. Larive, Comparison of GC-MS and NMR for metabolite profiling of rice subjected to submergence stress, J. Proteome Res. 12 (2013) 898–909.
[27] H. Dai, C. Xiao, H. Liu, F. Hao, H. Tang, Combined NMR and LC-DAD-MS analysis reveals comprehensive metabonomic variations for three phenotypic cultivars of salvia miltiorrhiza bunge, J. Proteome Res. 9 (2010) 1565–1578.
[28] K. Bingol, R. Brüschweiler, NMR/MS translator for the enhanced simultaneous analysis of metabolomics mixtures by NMR spectroscopy and mass spectrometry: application to human urine, J. Proteome Res. 14 (2015) 2642–2648.
[29] K. Bingol, L. Bruschweiler-Li, C. Yu, A. Somogyi, F. Zhang, R. Brüschweiler, Metabolomics beyond spectroscopic databases: a combined MS/NMR strategy for the rapid identification of new metabolites in complex mixtures, Anal. Chem. 87 (2015) 3864–3870.
[30] F. Tayyari, G.A. Gowda, H. Gu, D. Raftery, 15N-cholamine–a smart isotope tag for combining NMR- and MS-based metabolite profiling, Anal. Chem. 85 (2013) 8715–8721.
[31] J.M. Baker, J.L. Ward, M.H. Beale, Combined NMR and flow injection ESI-MS for Brassicaceae metabolomics, Meth. Mol. Biol. 860 (2012) 177–191.
[32] J. Prichystal, K.A. Schug, K. Lemr, J. Novak, V. Havlicek, Structural analysis of natural products, Anal. Chem. 88 (2016) 10338–10346.
[33] Z. Yang, Online hyphenated liquid chromatography-nuclear magnetic resonance spectroscopy-mass spectrometry for drug metabolite and nature product analysis, J. Pharm. Biomed. Anal. 40 (2006) 516–527.
[34] Z. Pan, D. Raftery, Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics, Anal. Bioanal. Chem. 387 (2007) 525–527.
[35] R. t'Kindt, R.A. Scheltema, A. Jankevics, K. Brunker, S. Rijal, J.C. Dujardin, R. Breitling, D.G. Watson, G.H. Coombs, S. Decuypere, Metabolomics to unveil and understand phenotypic diversity between pathogen populations, PLoS Negl. Trop. Dis. 4 (2010) e904.
[36] S. Halouska, B. Zhang, R. Gaupp, S. Lei, E. Snell, R.J. Fenton, R.G. Barletta, G.A. Somerville, R. Powers, Revisiting protocols for the NMR analysis of bacterial metabolomes, J. Integ. OMICS 2 (2013) 120–137.
[37] J.C. Copeland, L.J. Zehr, R.L. Cerny, R. Powers, The applicability of molecular descriptors for designing an electrospray ionization mass spectrometry compatible library for drug discovery, Comb. Chem. High Through. Screen. 15 (2012) 806–815.
[38] F.J. Moy, K. Haraki, D. Mobilio, G. Walker, K. Tabei, H. Tong, M.M. Siegel, R. Powers, MS/NMR: a structure-based approach for discovering protein ligands and for drug design by coupling size exclusion chromatography, mass spectrometry, and nuclear magnetic resonance spectroscopy, Anal. Chem. 73 (2001) 571–581.
[39] J.-P. Antignac, K. de Wasch, F. Monteau, H. De Brabander, F. Andre, B. Le Bizec, The ion suppression phenomenon in liquid chromatography-mass spectrometry and its consequences in the field of residue analysis, Anal. Chim. Acta 529 (2005) 129–136.
[40] T.O. Metz, J.S. Page, E.S. Baker, K. Tang, J. Ding, Y. Shen, R.D. Smith, High-resolution separations and improved ion production and transmission in metabolomics, TrAC Trends Anal. Chem. 27 (2008) 205–214.
[41] D.J. Crockford, E. Holmes, J.C. Lindon, R.S. Plumb, S. Zirah, S.J. Bruce, P. Rainville, C.L. Stumpf, J.K. Nicholson, Statistical Heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies, Anal. Chem. 78 (2006) 363–371.
[42] N.L. Kuehnbaum, P. Britz-McKibbin, New advances in separation science for metabolomics: resolving chemical diversity in a post-genomic era, Chem. Rev. 113 (2013) 2437–2468.
[43] D.B. Kell, Metabolomics and systems biology: making sense of the soup, Curr. Opin. Microbiol. 7 (2004) 296–307.
[44] A.B. Canelas, A. ten Pierick, C. Ras, R.M. Seifar, J.C. van Dam, W.M. van Gulik, J. J. Heijnen, Quantitative evaluation of intracellular metabolite extraction techniques for yeast metabolomics, Anal. Chem. 81 (2009) 7379–7389.
[45] H. Kanani, P.K. Chrysanthopoulos, M.I. Klapa, Standardizing GC-MS metabolomics, J. Chromatogr. B: Anal. Technol. Biomed., Life Sci. 871 (2008) 191–201.
[46] F. Xu, L. Zou, C.N. Ong, Multiorigination of chromatographic peaks in derivatized GC/MS metabolomics: a confounder that influences metabolic pathway interpretation, J. Proteome Res. 8 (2009) 5657–5665.
[47] P.J. Taylor, Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography-electrospray-tandem mass spectrometry, Clin. Biochem. 38 (2005) 328–334.

[48] J. Kopka, Current challenges and developments in GC-MS based metabolite profiling technology, J. Biotechnol. 124 (2006) 312–322.
[49] D. Marion, P.C. Driscoll, L.E. Kay, P.T. Wingfield, A. Bax, A.M. Gronenborn, G.M. Clore, Overcoming the overlap problem in the assignment of ^{1}H NMR spectra of larger proteins by use of three-dimensional heteronuclear ^{1}H–^{15}N Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta, Biochemistry 28 (1989) 6150–6156.
[50] E.A. Mahrous, M.A. Farag, Two dimensional NMR spectroscopic approaches for exploring plant metabolome: a review, J. Adv. Res. 6 (2015) 3–15.
[51] K. Bingol, L. Brüschweilerr-Li, D. Li, B. Zhang, M. Xie, R. Brüschweiler, Emerging new strategies for successful metabolite identification in metabolomics, Bioanalysis 8 (2016) 557–573.
[52] Z. Lei, D.V. Huhman, L.W. Sumner, Mass spectrometry strategies in metabolomics, J. Biol. Chem. 286 (2011) 25435–25442.
[53] Y. Iwasaki, T. Sawada, K. Hatayama, A. Ohyagi, Y. Tsukuda, K. Namekawa, R. Ito, K. Saito, H. Nakazawa, Separation technique for the determination of highly polar metabolites in biological samples, Metabolites 2 (2012) 496–515.
[54] N.B. Cech, C.G. Enke, Practical implications of some recent studies in electrospray ionization fundamentals, Mass Spectrom. Rev. 20 (2001) 362–387.
[55] F. Tugizimana, P.A. Steenkamp, L.A. Piater, I.A. Dubery, Multi-platform metabolomic analyses of ergosterol-induced dynamic changes in Nicotiana tabacum cells, PLoS One 9 (2014) e87846.
[56] Y.S. Ling, H.J. Liang, M.H. Chung, M.H. Lin, C.Y. Lin, NMR- and MS-based metabolomics: various organ responses following naphthalene intervention, Mol. Biosyst. 10 (2014) 1918–1931.
[57] D.D. Marshall, S. Lei, B. Worley, Y. Huang, A. Garcia-Garcia, R. Franco, E.D. Dodds, R. Powers, Combining DI-ESI-MS and NMR datasets for metabolic profiling, Metabolomics 11 (2015) 391–402.
[58] K. Bingol, R. Brüschweiler, Two elephants in the room, Curr. Opin. Clin. Nut. Metab. Care 18 (2015) 471–477.
[59] Y. Lin, S. Schiavo, J. Orjala, P. Vouros, R. Kautz, Microscale LC-MS-NMR platform applied to the identification of active cyanobacterial metabolites, Anal. Chem. 80 (2008) 8045–8054.
[60] O. Corcoran, M. Spraul, LC–NMR–MS in drug discovery, Drug Discov. Today 8 (2003) 624–631.
[61] C.H. Johnson, J. Ivanisevic, G. Siuzdak, Metabolomics: beyond biomarkers and towards mechanisms, Nat. Rev. Mol. Cell Biol. 17 (2016) 451–459.
[62] S.S. Bird, D.P. Sheldon, R.M. Gathungu, P. Vouros, R. Kautz, W.R. Matson, B.S. Kristal, Structural characterization of plasma metabolites detected via LC-electrochemical coulometric array using LC-UV fractionation, MS, and NMR, Anal. Chem. 84 (2012) 9889–9898.
[63] A. Alonso, S. Marsal, A. Julia, Analytical methods in untargeted metabolomics: state of the art in 2015, Front. Bioeng. Biotechnol. 3 (2015) 23.
[64] W.B. Dunn, D.I. Ellis, Metabolomics: current analytical platforms and methodologies, TrAC Trends Anal. Chem. 24 (2005) 285–294.
[65] B. Worley, R. Powers, PCA as a predictor of OPLS-DA model reliability, Curr. Metabolom. 4 (2016) 97–103.
[66] R.G. Brereton, A short history of chemometrics: a personal view, J. Chemom. 28 (2014) 749–760.
[67] B. Worley, R. Powers, MVAPACK: a complete data handling package for NMR metabolomics, ACS Chem. Biol. 9 (2014) 1138–1144.
[68] B. Worley, R. Powers, A sequential algorithm for multiblock orthogonal projections to latent structures, Chemometr. Intell. Lab. Syst. 149 (2015) 33–39.
[69] H. Chen, Z. Pan, N. Talaty, D. Raftery, R.G. Cooks, Combining desorption electrospray ionization mass spectrometry and nuclear magnetic resonance for differential metabolomics without sample preparation, Rapid Commun. Mass Spectrom. 20 (2006) 1577–1584.
[70] H. Gu, Z. Pan, B. Xi, V. Asiago, B. Musselman, D. Raftery, Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: application to the detection of breast cancer, Anal. Chim. Acta 686 (2011) 57–63.
[71] S. Lei, L. Zavala-Flores, A. Garcia-Garcia, R. Nandakumar, Y. Huang, N. Madayiputhiya, R.C. Stanton, E.D. Dodds, R. Powers, R. Franco, Alterations in energy/redox metabolism induced by mitochondrial and environmental toxins: a specific role for glucose-6-phosphate-dehydrogenase and the pentose phosphate pathway in paraquat toxicity, ACS Chem. Biol. 9 (2014) 2032–2048.
[72] A. Anandhan, S. Lei, R. Levytskyy, A. Pappa, M.I. Panayiotidis, R.L. Cerny, O. Kalimonchuk, R. Powers, R. Franco, Glucose Metabolism and AMPK signaling regulate dopaminergic cell death induced by gene (α-synuclein)-environment (paraquat) interactions, Mol. Neurobiol. (2016), http://dx.doi.org/10.1007/s12035-016-9906-2.
[73] R. Powers, NMR metabolomics and drug discovery, Magn. Reson. Chem. 47 (2009) S2–S11.
[74] M.E. Bollard, E.G. Stanley, J.C. Lindon, J.K. Nicholson, E. Holmes, NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition, NMR Biomed. 18 (2005) 143–162.
[75] W.B. Dunn, D.I. Broadhurst, H.J. Atherton, R. Goodacre, J.L. Griffin, Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy, Chem. Soc. Rev. 40 (2011) 387–426.
[76] J.K. Nicholson, J.C. Lindon, E. Holmes, "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data, Xenobiotica 29 (1999) 1181–1189.
[77] D.S. Wishart, Quantitative metabolomics using NMR, TrAC, Trends Anal. Chem. 27 (2008) 228–237.
[78] K. Chughtai, R.M.A. Heeren, Mass spectrometric imaging for biomedical tissue analysis, Chem. Rev. 110 (2010) 3237–3277.
[79] M. Katajamaa, M. Oresic, Data processing for mass spectrometry-based metabolomics, J. Chromatogr. A 1158 (2007) 318–328.
[80] S.G. Villas-Boas, S. Mas, M. Aakesson, J. Smedsgaard, J. Nielsen, Mass spectrometry in metabolome analysis, Mass Spectrom. Rev. 24 (2005) 613–646.
[81] A.C. Dona, B. Jimenez, H. Schafer, E. Humpfer, M. Spraul, M.R. Lewis, J.T.M. Pearce, E. Holmes, J.C. Lindon, J.K. Nicholson, Precision high-throughput proton nmr spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping, Anal. Chem. 86 (2014) 9887–9894.
[82] A.-H. Emwas, C. Luchinat, P. Turano, L. Tenori, R. Roy, R.M. Salek, D. Ryan, J.S. Merzaban, R. Kaddurah-Daouk, A.C. Zeri, G.A. Nagana Gowda, D. Raftery, Y. Wang, L. Brennan, D.S. Wishart, Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review, Metabolomics 11 (2015) 872–894.
[83] D. Troyer, Biorepository standards and protocols for collecting, processing, and storing human tissues, Meth. Mol. Biol. (Totowa, NJ, US) 441 (2008) 193–220.
[84] S.G. Villas-Boas, J. Hoejer-Pedersen, M. Aakesson, J. Smedsgaard, J. Nielsen, Global metabolite analysis of yeast: evaluation of sample preparation methods, Yeast 22 (2005) 1155–1169.
[85] W.B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J.D. Knowles, A. Halsall, J.N. Haselden, A.W. Nicholls, I.D. Wilson, D.B. Kell, R. Goodacre, Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry, Nat. Protoc. 6 (2011) 1060–1083.
[86] A.-H. Zhang, H. Sun, S. Qiu, X.-J. Wang, NMR-based metabolomics coupled with pattern recognition methods in biomarker discovery and disease diagnosis, Magn. Reson. Chem. 51 (2013) 549–556.
[87] J.A. Westerhuis, H.C.J. Hoefsloot, S. Smit, D.J. Vis, A.K. Smilde, E.J.J. van Velzen, J.P.M. van Duijnhoven, F.A. van Dorsten, Assessment of PLSDA cross validation, Metabolomics 4 (2008) 81–89.
[88] K. Kjeldahl, R. Bro, Some common misunderstandings in chemometrics, J. Chemom. 24 (2010) 558–564.
[89] L. Eriksson, J. Trygg, S. Wold, CV-ANOVA for significance testing of PLS and OPLS models, J. Chemom. 22 (2008) 594–600.
[90] T.A. Lasko, J.G. Bhagwat, K.H. Zou, L. Ohno-Machado, The use of receiver operating characteristic curves in biomedical informatics, J. Biomed. Inf. 38 (2005) 404–415.
[91] J. Xia, D.S. Wishart, MetPA: a web-based metabolomics tool for pathway analysis and visualization, Bioinformatics 26 (2010) 2342–2344.
[92] J.J. Heijnen, Impact of thermodynamic principles in systems biology, Adv. Biochem. Eng./Biotechnol. 121 (2010) 139–162.
[93] S. Arrivault, M. Guenther, A. Ivakov, R. Feil, D. Vosloh, J.T. van Dongen, R. Sulpice, M. Stitt, Use of reverse-phase liquid chromatography, linked to tandem mass spectrometry, to profile the Calvin cycle and other metabolic intermediates in Arabidopsis rosettes at different carbon dioxide concentrations, Plant J. 59 (2009) 826–839.
[94] M. Breier, S. Wahl, C. Prehn, M. Fugmann, U. Ferrari, M. Weise, F. Banning, J. Seissler, H. Grallert, J. Adamski, A. Lechner, Targeted metabolomics identifies reliable and stable metabolites in human serum and plasma samples, PLoS One 9 (2014), e89728/89721-e89728/89711, 89711 pp.
[95] R.A. van den Berg, H.C.J. Hoefsloot, J.A. Westerhuis, A.K. Smilde, M.J. van der Werf, Centering, scaling, and transformations: improving the biological information content of metabolomics data, BMC Genom. 7 (2006) 142.
[96] A. Craig, O. Cloarec, E. Holmes, J.K. Nicholson, J.C. Lindon, Scaling and normalization effects in NMR spectroscopic metabonomic data sets, Anal. Chem. 78 (2006) 2262–2267.
[97] T.N. Vu, K. Laukens, Getting your peaks in line: a review of alignment methods for NMR spectral data, Metabolites 3 (2013) 259–276.
[98] C.L. Andrews, C.P. Yu, E. Yang, P. Vouros, Improved liquid chromatography-mass spectrometry performance in quantitative analysis using a nanosplitter interface, J. Chromatogr. A 1053 (2004) 151–159.
[99] A.H. Emwas, R. Roy, R.T. McKay, D. Ryan, L. Brennan, L. Tenori, C. Luchinat, X. Gao, A.C. Zeri, G.A. Gowda, D. Raftery, C. Steinbeck, R.M. Salek, D.S. Wishart, Recommendations and standardization of biomarker quantification using NMR-based metabolomics with particular focus on urinary analysis, J. Proteome Res. 15 (2016) 360–373.
[100] E. Appiah-Amponsah, K. Owusu-Sarfo, G.A. Gowda, T. Ye, D. Raftery, Combining hydrophilic interaction chromatography (HILIC) and isotope tagging for off-line LC-NMR applications in metabolite analysis, Metabolites 3 (2013) 575–591.
[101] A. Beltran, M. Suarez, M.A. Rodríguez, M. Vinaixa, S. Samino, L. Arola, X. Correig, O. Yanes, Assessment of compatibility between extraction methods for NMR- and LC/MS-based metabolomics, Anal. Chem. 84 (2012) 5838–5844.
[102] A.L. Harvey, R. Edrada-Ebel, R.J. Quinn, The re-emergence of natural products for drug discovery in the genomics era, Nat. Rev. Drug Discovery 14 (2015) 111–129.
[103] B. David, J.-L. Wolfender, D.A. Dias, The pharmaceutical industry and natural products: historical status and new trends, Phytochem. Rev. 14 (2014) 299–315.

[104] J.J. Kellogg, D.A. Todd, J.M. Egan, H.A. Raja, N.H. Oberlies, O.M. Kvalheim, N.B. Cech, Biochemometrics for natural products research: comparison of data analysis approaches and application to identification of bioactive compounds, J. Nat. Prod. 79 (2016) 376–386.

[105] F. Bucar, A. Wube, M. Schmid, Natural product isolation–how to get from biological material to pure compounds, Nat. Prod. Rep. 30 (2013) 525–545.

[106] A. Krunic, J. Orjala, Application of high-field NMR spectroscopy for characterization and quantitation of submilligram quantities of isolated natural products, Magn. Reson. Chem. 53 (2015) 1043–1050.

[107] T.F. Molinski, Nanomole-scale natural products discovery, Curr. Opin. Drug Disc. Dev. 12 (2009) 197–206.

[108] D.L. Olson, J.A. Norcross, M. O'Neil-Johnson, P.F. Molitor, D.J. Detlefsen, A.G. Wilson, T.L. Peck, Microflow NMR: concepts and capabilities, Anal. Chem. 76 (2004) 2966–2974.

[109] K.C. Nicolaou, S.A. Snyder, Chasing molecules that were never there: misassigned natural products and the role of chemical synthesis in modern structure elucidation, Angew. Chem., Int. Ed. 44 (2005) 1012–1044.

[110] K. Biemann, High-resolution mass spectrometry of natural products, Pure Appl. Chem. 9 (1964) 95–118.

[111] K.-W. Cheng, C.-C. Wong, M. Wang, Q.-Y. He, F. Chen, Identification and characterization of molecular targets of natural products by mass spectrometry, Mass Spectrom. Rev. 29 (2010) 126–155.

[112] R.G. Cooks, D.R. Ifa, G. Sharma, F.K. Tadjimukhamedov, Z. Ouyang, Perspectives and retrospectives in mass spectrometry: one view, Eur. J. Mass Spectrom. 16 (2010) 283–300.

[113] K. Biemann, Structure determination of natural products by mass spectrometry, Annu. Rev. Anal. Chem. 8 (2015) 1–19.

[114] S.P. Gaudencio, F. Pereira, Dereplication: racing to speed up the natural products discovery process, Nat. Prod. Rep. 32 (2015) 779–810.

[115] I.F. Duarte, M. Godejohann, U. Braumann, M. Spraul, A.M. Gil, Application of NMR spectroscopy and LC-NMR/MS to the identification of carbohydrates in beer, J. Agric. Food Chem. 51 (2003) 4847–4852.

[116] J.P. Shockcor, S.E. Unger, P. Savina, J.K. Nicholson, J.C. Lindon, Application of directly coupled LC-NMR-MS to the structural elucidation of metabolites of the HIV-1 reverse-transcriptase inhibitor BW935U83, J. Chromatogr. B Biomed. Sci. Appl. 748 (2000) 269–279.

[117] M. Spraul, A.S. Freund, R.E. Nast, R.S. Withers, W.E. Maas, O. Corcoran, Advancing NMR sensitivity for LC-NMR-MS using a cryoflow probe: application to the analysis of acetaminophen metabolites in urine, Anal. Chem. 75 (2003) 1536–1541.

[118] Y. Iwasaki, Y. Nakano, K. Mochizuki, M. Nomoto, Y. Takahashi, R. Ito, K. Saito, H. Nakazawa, A new strategy for ionization enhancement by derivatization for mass spectrometry, J. Chromatogr. B: Anal. Technol. Biomed., Life Sci. 879 (2011) 1159–1165.

[119] T. Santa, Derivatization reagents in liquid chromatography/electrospray ionization tandem mass spectrometry, Biomed. Chromatogr. 25 (2011) 1–10.

[120] J.M. Halket, D. Waterman, A.M. Przyborowska, R.K.P. Patel, P.D. Fraser, P.M. Bramley, Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS, J. Exp. Bot. 56 (2005) 219–243.

[121] J.M. Halket, V.G. Zaikin, Derivatization in mass spectrometry-1. Silylation, Eur. J. Mass Spectrom. 9 (2003) 1–21.

[122] M. Brattoli, E. Cisternino, P.R. Dambruoso, G. de Gennaro, P. Giungato, A. Mazzone, J. Palmisani, M. Tutino, Gas chromatography analysis with olfactometric detection (GC-O) as a useful methodology for chemical characterization of odorous compounds, Sensors (Basel) 13 (2013) 16759–16800.

[123] G.A. Gowda, F. Tayyari, T. Ye, Y. Suryani, S. Wei, N. Shanaiah, D. Raftery, Quantitative analysis of blood plasma metabolites using isotope enhanced NMR methods, Anal. Chem. 82 (2010) 8983–8990.

[124] A.N. Lane, T.W. Fan, Z. Xie, H.N. Moseley, R.M. Higashi, Isotopomer analysis of lipid biosynthesis by high resolution mass spectrometry and NMR, Anal. Chim Acta 651 (2009) 201–208.

[125] A.N. Lane, T.W.M. Fan, M. Bousamra, R.M. Higashi, J. Yan, D.M. Miller, Stable isotope-resolved metabolomics (SIRM) in cancer res. with clinical application to nonsmall cell lung cancer, OMICS 15 (2011) 173–182.

[126] T.W.M. Fan, A.N. Lane, Applications of NMR spectroscopy to systems biochemistry, Prog. Nucl. Magn. Reson. Spectrosc. 92–93 (2016) 18–53.

[127] T.W.M. Fan, A.N. Lane, R.M. Higashi, M.A. Farag, H. Gao, M. Bousamra, D.M. Miller, Altered regulation of metabolic pathways in human lung cancer discerned by (13)C stable isotope-resolved metabolomics (SIRM), Mol. Can. 8 (2009) 41.

[128] T.W.M. Fan, P. Yuan, A.N. Lane, R.M. Higashi, Y. Wang, A.B. Hamidi, R. Zhou, X. Guitart, G. Chen, H.K. Manji, R. Kaddurah-Daouk, Stable isotope-resolved metabolomic analysis of lithium effects on glial-neuronal metabolism and interactions, Metabolomics 6 (2010) 165–179.

[129] R.M. Higashi, T.W.M. Fan, P.K. Lorkiewicz, H.N.B. Moseley, A.N. Lane, Stable isotope-labeled tracers for metabolic pathway elucidation by GC-MS and FT-MS, 1198 (2014) 147-167.

[130] A.N. Lane, R.M. Higashi, T.W.M. Fan, Preclinical models for interrogating drug action in human cancers using stable isotope resolved metabolomics (SIRM), Metabolomics 12 (2016) 1–15.

[131] T.W.M. Fan, A.N. Lane, R.M. Higashi, J. Yan, Stable isotope resolved metabolomics of lung cancer in a SCID mouse model, Metabolomics 7 (2011) 257–269.

[132] T.W.M. Fan, P.K. Lorkiewicz, K. Sellers, H.N.B. Moseley, R.M. Higashi, A.N. Lane, Stable isotope-resolved metabolomics and applications for drug development, Pharmacol. Ther. 133 (2012) 366–391.

[133] H.N. Moseley, Correcting for the effects of natural abundance in stable isotope resolved metabolomics experiments involving ultra-high resolution mass spectrometry, BMC Bioinf. 11 (2010) 139.

[134] W. Callebaut, Scientific perspectivism: a philosopher of science's response to the challenge of big data biology, Stud. Hist. Philos. Biol. Biomed. Sci. 43 (2012) 69–80.

[135] K. Dolinski, O.G. Troyanskaya, Implications of big data for cell biology, Mol. Biol. Cell. 26 (2015) 2575–2578.

[136] V. Gligorijevic, N. Malod-Dognin, N. Przulj, Integrative methods for analyzing big data in precision medicine, Proteomics 16 (2016) 741–758.

[137] J.D. Tenenbaum, S.-A. Sansone, M. Haendel, A sea of standards for omics data: sink or swim?, J Am. Med. Inf. Assoc. 21 (2014) 200–203.

[138] I. Anagnostopoulos, S. Zeadally, E. Exposito, Handling big data: research challenges and future directions, J. Supercomput. 72 (2016) 1494–1516.

[139] S.L. Robinette, F. Zhang, L. Bruschweiler-Li, R. Bruschweiler, Web server based complex mixture analysis by NMR, Anal. Chem. 80 (2008) 3606–3611.

[140] H.E. Pence, A. Williams, ChemSpider: an online chemical information resource, J. Chem. Educ. 87 (2010) 1123–1124.

[141] A.M.D. Livera, M. Sysi-Aho, L. Jacob, J.A. Gagnon-Bartsch, S. Castillo, J.A. Simpson, T.P. Speed, Statistical methods for handling unwanted variation in metabolomics data, Anal. Chem. 87 (2015) 3606–3615.

[142] S. Wold, W. S., S. Hellbern, T. Lundstedt, M. Siostrom, PLS modeling with latent variables in two or more dimensions, in: Proceedings PLS Model Building, Theory and Applications Symposium Frankfurt am Main September 23–25, 1987.

[143] A.K. Smilde, J.A. Westerhuis, S. de Jong, A framework for sequential multiblock component methods, J. Chemom. 17 (2003) 323–337.

[144] J.A. Westerhuis, T. Kourti, J.F. Macgregor, Analysis of multiblock and hierarchical PCA and PLS models, J. Chemom. 12 (1998) 301–321.

[145] D.E. Befroy, R.J. Perry, N. Jain, S. Dufour, G.W. Cline, J.K. Trimmer, J. Brosnan, D.L. Rothman, K.F. Petersen, G.I. Shulman, Direct assessment of hepatic mitochondrial oxidative and anaplerotic fluxes in humans using dynamic ^{13}C magnetic resonance spectroscopy, Nat. Med. 20 (2014) 98–102.

Glossary of Abbreviations

1D: One-Dimensional

2D: Two-Dimensional

3D: Three-Dimensional

6-OHDA: 6-hydroxydopamine

ANOVA: ANalysis Of Variance

BNMI: Bruker NMR-Mass Spectrometry Interface

CE: Capillary Electrophoresis

COLMAR: Complex Mixture Analysis by NMR

CV-ANOVA: ANOVA of the Cross-Validated residuals

ddH$_2$O: double-distilled water

DI-ESI-MS: Direct Infusion ElectroSpray Ionisation Mass Spectrometry

ESI: ElectroSpray Ionisation

GC: Gas Chromatograph

GPL: GlyceroPhosphoLipids

HMBC: Heteronuclear Multiple-Bond Correlation

HSQC: Heteronuclear Single Quantum Coherence spectroscopy

FT-ICR-MS: Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

LC: Liquid Chromatograph

LysoPC (16:0): Lysophospholipid (16:0)

MB-PCA: Multi-Block-Principal Component Analysis

MB-PLS: Multi-Block-Projections to Latent Structures

MetPA: METabolomics Pathway Analysis

MPP$^+$: 1-methyl-4-phenylpyridinium

MS: Mass Spectrometry

MS^n: Multiple-stage mass spectrometry where n is the number of product ion stages (n = 2,3, *etc.*)

m/z: mass to charge ratio

NADPH: Nicotinamide Adenine Dinucleotide Phosphate

NMR: Nuclear Magnetic Resonance

OPLS: Orthogonal Projections to Latent Structures

OPLS-DA: Orthogonal Projections to Latent Structures- Discriminate Analysis

PC (34:1): PhosphatidylCholine 34:1

PC: Principal Component

PC1: first Principal Component

PCA: Principal Component Analysis

pKa: negative logarithm of an acid dissociation constant

PLS: Projections to Latent Structures

PPP: Pentose Phosphate Pathway

PRA: Pattern Recognition Approaches

ROC: Receiver Operating Characteristic

RSD: Relative Standard Deviation

SIRM: Stable Isotope-Resolved Metabolomics

TCA cycle: TriCarboxylic Acid cycle

TMSP-d$_4$: 3-(TriMethylSilyl)Propionic-2,2,3,3-d$_4$

TOCSY: Total Correlation SpectroscopY

XCMS: various forms (X) of Chromatography Mass Spectrometry