

RESEARCH ARTICLE

[View Article Online](#)
[View Journal](#) | [View Issue](#)

The robust NMR toolbox for metabolomics†

Cite this: *Mol. Omics*, 2021,
17, 719Kousik Chandra,^a Samah Al-Harthi,^a Fatimah Almulhim,^a
Abdul-Hamid Emwas,^b Łukasz Jaremko^{*a} and Mariusz Jaremko^{†a}Received 14th April 2021,
Accepted 25th May 2021

DOI: 10.1039/d1mo00118c

rsc.li/molomics

Here, we implemented and validated a suite of selective and non-selective CPMG-filtered 1D and 2D TOCSY/HSQC experiments for metabolomics research. They facilitated the unambiguous identification of metabolites embedded in broad lipid and protein signals. The 2D spectra improved non-targeted analysis by removing the background broad signals of macromolecules.

NMR is the most widely used spectroscopy technique in metabolomics after mass spectrometry. NMR advantages include straightforward and non-invasive sample preparation, standardized signal quantification, the ability to monitor several metabolites in a single experiment, *etc.* However, the complex bio-fluid spectra often exhibit an extensive signal dynamic range accompanied by severe overlap between metabolites and macromolecules. Thus, an unambiguous assignment often becomes tedious or impossible. The precise identification of chemical moieties, the spin-system connectivity, and finally, the molecules' chemical identity and abundance are essential for analyzing mixtures of metabolites.¹ The assignment is the critical step for metabolite identification, which requires the recording of multiple datasets. This is accompanied by high-throughput data from numerous samples of bio-fluids, followed by several multivariate data analyses.² The workflow's precise aim becomes identifying the biomarkers by monitoring specific changes in the bio-fluid metabolite composition.

The advantage of NMR metabolomics is the simultaneous monitoring of the entire sample content, including metabolites and other entities like macromolecules. However, it is also often a limitation. Real samples are composed of hundreds or thousands of different compounds with a very high dynamic range of concentrations and signal shape. As a result, the severe overlap of signals from a few highly concentrated components dominates multivariate data analysis, such as principal component analysis (PCA).³ However, it is essential to also

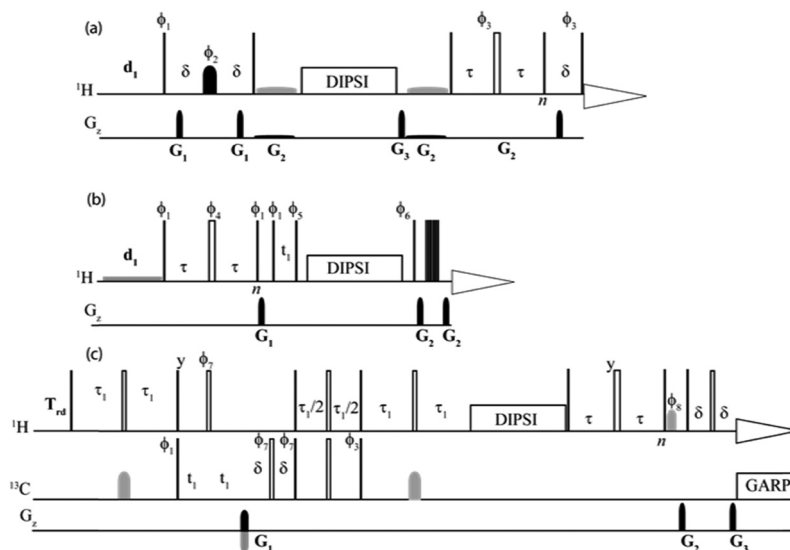
study the minor components as they dominate the biochemical and physiological properties and can be utilized as specific biomarkers of a disease or state.⁴ Targeted analysis explicitly monitors a predefined set of compounds and biomarkers.⁵ The greatest challenge in the targeted analysis is the presence of large macromolecules like proteins, lipids, and membranes that cover the signals from the small metabolites.⁶ For example, it is highly challenging in human blood serum samples to identify and quantify NMR signals arising from branched-chain amino acids (BCAAs) like valine, leucine, and isoleucine (resonating at 0.92–0.99 ppm). This is because the lower signals arising from BCAAs are substantially overlapped by the broad lipid signal resonating at 0.89 ppm ($-\text{CH}_3$ groups).² Since the BCAA signals bear some lipid components, relaxation-based editing has been commonly used to precisely observe the slowly relaxing molecules in bio-fluid.^{7,8}

To address this challenge, we combine and implement total correlation spectroscopy (TOCSY)⁹-based magnetization transfer, transverse relaxation editing within a minimum set of pulse programs and subsequently validate this set for robust targeted and non-targeted analysis of bovine serum. There has been phenomenal advancement regarding the fast and efficient acquisition of homonuclear experiments for labile sites in proteins¹⁰ based on the HADAMARD principle.¹¹ Recently significant advancement was made feasible by the application of PANSY (Parallel NMR Spectroscopy) experiments to obtain sufficient spectral information for the structure elucidation of small organic molecules.¹² Diffusion ordered (DO) NMR spectroscopy has combined with statistical total correlation spectroscopy (STOCSY)¹³ for complex bio-mixture analysis.¹⁴ Here we take an approach that combines the previously developed advanced features like CPMG or TOCSY and ¹³C dispersion in the indirect dimension *etc.*, and the main aim is to present the minimum number of experiments to encompass

^a Biological and Environmental Science and Engineering (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. E-mail: Mariusz.jaremko@kaust.edu.sa

^b Core Laboratories, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1mo00118c



structure determination in TOCSY optimized mixture elucidation (TOOMIXED).²⁸ We have combined the previous developments here. 1D selective ^1H TOCSY employs a selective 180° shape pulse (Reburp)²⁹ sandwiched between two z gradients to achieve selective excitation of any peak of the desired target molecule. Subsequently, the TOCSY spinlock is applied with a z -filter. For spinlock, DIPSI was used. The shape pulse was flanked by Gaussian-shaped pulsed z -field gradients of 1 ms duration each. A Reburp-shaped selective 180° inversion pulse with durations of 300 ms was the most effective for selective excitations among most overlapped regions. We also used 80 ms Reburp; however, in highly overlapped methyl regions, the 300 ms selective pulse worked better for bovine serum (Fig. S1, ESI†). The Reburp was chosen for its better off-resonance effect and higher sensitivity^{30,31} compared to the normally used Gauss pulse (Fig. S2, ESI†). It is worth mentioning that as we chose a more selective pulse, the sensitivity decreased (Fig. S3, ESI†). However, for a one-dimensional experiment, this does not matter. We get good sensitivity even with a 300 ms selective pulse with recycle delay of 5 seconds in 2 minutes (Fig. S4, ESI†). Since selective 1D ^1H TOCSY is generally used for targeted analysis, it is more important to achieve clean spectra for the chosen metabolites. The optimal TOCSY mixing time was found to be 80 ms. Simultaneously, a swept-frequency adiabatic chirp 180° pulse³² and the low gradient were applied before and after spinlock for suppressing zero quantum coherences.³³ Because of the applied gradient along the z -axis, zero quantum coherences evolve and refocus at $2\alpha T_f$ where α varies between 0 and 1 along the z -axis of the tube height, and T_f is the duration of the adiabatic pulse. For the time $(1 - 2\alpha)T_f$, zero quantum coherences evolve and acquire different phases along the tube and get canceled. Subsequently, the magnetization is brought into the transverse plane by applying a 90° pulse, and the CPMG spin-echo sequence is employed for the T_2 editing of large bio macromolecules. The CPMG spin-echo sequence (τ - 180° - τ) consists of spin-echo time (τ , half of the delay between two successive 180° pulses), one hard 180° pulse, and several repetitive loops (n). We kept the spin-echo time as 1 ms. A total of 120 ms T_2 -relaxation delay is sufficient to attenuate the resonances from large macromolecules and small molecules bound to them due to large differences in their respective tumbling times. It is crucial to note that keeping a longer relaxation time decreases the sensitivity and the appropriate spacing between the refocusing pulses necessary for probe safety and to avoid sample heating. For a selective 1D ^1H experiment, the sensitivity is not a problem as this is less time-consuming, and a longer T_2 -filter of 240 ms can easily be used. Usage of CPMG is especially required for less selective pulses like the conventionally used 40 ms Reburp (Fig. S5, ESI†). However, for multidimensional experiments, the experimental time increases, and appropriate optimization should be taken for minimal T_2 -filter and selective pulse duration.

We implemented 1D selective ^1H CPMG TOCSY on bovine serum samples dissolved in D_2O . As a demonstrative example,

we chose the methyl region, where a broad lipid peak and protein resonance dominates, severely crowded sugar regions, and one less crowded region (marked as a–c in Fig. 2, respectively). We could filter out the distinct signals of valine, glucose, and succinate from these regions, as shown in Fig. 2. The corresponding offsets are mentioned within the panel (a1–c1). The chemical shifts along with peak pattern and peak intensities can easily be matched in the standard metabolite library for each candidate. We found excellent agreement with the metabolite libraries provided by Chenomx,³⁴ as shown in Fig. S6 (ESI†). For the highly congested sugar regions, where chemical shifts differ very minutely, the chances of exciting multiple resonances are very high. For example, when the offset was chosen at 3.2701 ppm for one of the glucose peaks, the resultant spectra did not match the library's standard glucose peak patterns, however, shifting the offset to 3.4233 ppm unambiguously assigned the resonances of glucose (see Fig. S7, ESI†). In the less crowded region, things were more straightforward, as depicted for succinate resonance. In fact, we need to find one isolated peak for the given compound. For example, the offset at 1.045 ppm gives a clear pattern of valine only. However, we find that shifting the offset to 3.623 or 2.275 ppm excites multiple peaks. Still, based on the TOCSY pattern, these chemical shifts can be confirmed (Fig. S8, ESI†). Things are much simpler for isolated resonances like lactate, where the shifting of the offset between 4.119 and 1.333 ppm yields identical spectra as here both the peaks are in the isolated region. But the advantage is, still, one can assign a peak in the crowded region based on the TOCSY pattern. Additionally, adding CPMG helped in the complete removal of broad peaks of macromolecules. This technique retained all the advantages of selective TOCSY and proved highly robust in all the scenarios demonstrated above.

Our main aim was to provide a complete set of experiments for an unambiguous assignment for targeted and non-targeted analysis. For the second case, it is desirable to observe all the resonances in a single spectrum. In that case, in the one-dimensional NMR of complex biofluids, the situation becomes limited due to signal overlap from various metabolites. A computational approach was proposed from an analysis perspective where spectral deconvolution fits individual peaks in crowded spectra.³⁵ Commercially available software like Chenomx³⁴ also provides the option to fit the metabolites' peaks in the libraries to adjust the effects of temperature and pH on chemical shifts. However, from the data acquisition approach, using multidimensional NMR techniques is most suitable as the signal could be spread in two to three dimensions, which increases the specificity of the signals. Since biological samples of metabolomics rely on natural abundance, 2D NMR is feasible here.³⁶ This provides the opportunity to use a nucleus other than ^1H , with hetero-nuclei like ^{13}C being the most obvious alternative. Here we implanted CPMG on both non-selective homonuclear and heteronuclear TOCSY. On the homonuclear 2D ^1H - ^1H CPMG TOCSY, data were collected in a phase-sensitive manner with coherence selection achieved using an Echo/Antiecho gradient pair.³⁷ The water suppression was achieved using a 3-9-19 pulse sequence along with the gradient pair before detection.^{38,39} We also employed excitation



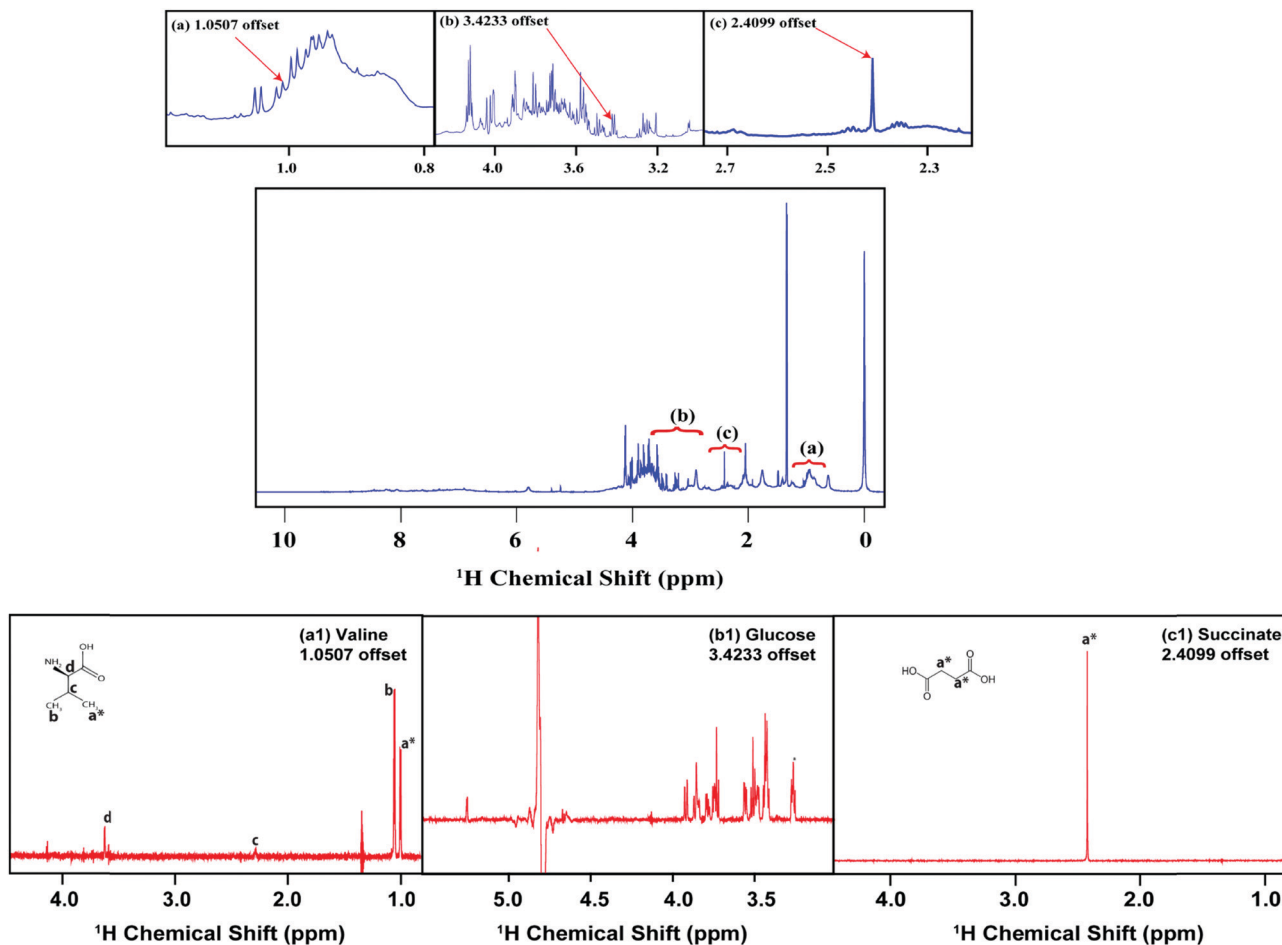


Fig. 2 1D ^1H spectra bovine serum samples dissolved in D_2O . Three different regions are indicated: (a) a region dominated by a broad lipid peak and protein resonance, (b) severely crowded sugar regions, and (c) a less crowded region. The regions are expanded in the sub panels above. The offset frequencies used for selective excitation in 1D selective ^1H CPMG TOCSY are denoted within each panel along with the peak position highlighted with a red arrow. The 1D selective ^1H CPMG TOCSY spectra with different offset frequencies are denoted within each panel of (a1), (b1) and (c1). Distinct signals of valine (a1), glucose (b1), and succinate (c1) from regions (a), (b) and (c) are shown respectively.

sculpting for water suppression. 3-9-19 methods work better close to the water signal at 4.5 ppm. However, some off-resonance peaks around 2.2 ppm showed better sensitivity upon excitation sculpting (Fig. S9, ESI †).

Additionally, pre-saturation of the water signal was employed during the relaxation delay (Fig. 1b). T_2 editing of large bio-macromolecules was achieved by applying a CPMG spin-echo sequence directly after the first pulse, bringing magnetization in the transverse plane. As discussed earlier, the optimized value of 120 ms T_2 delay along 80 ms mixing time was used for optimal sensitivity. Fig. S5 (ESI †) demonstrates that CPMG TOCSY retains the same peaks as 2D TOCSY. The overlay figure shows that the broad signal from lipid and protein molecules gets reduced in the highlighted regions. Overall the spectra become much cleaner, as evident from the highlighted region of the overlay of the trace taken at 3.8 ppm (Fig. S10(d), ESI †), and signatures from metabolites become more prominent. Since moving to multidimensional NMR increases the measurement time, certain methodological developments have been proposed. Here we employed non-uniform sampling,

which yielded excellent resolution (Fig. S11, ESI †) in half of the measurement time.⁴⁰ In fact, we can go down easily to 25% NUS as most of the peaks are retained until that point (Fig. S12, ESI †). Below this, with 12.5% data, some artefacts start coming (Fig. S12, ESI †). Earlier selective optimized flip angle short transient (SOFAST) techniques have been implemented for experimental time reduction.³⁶ However, for the non-targeted analysis TOCSY experiments performed here, this was not an option. We also achieved ^{13}C based 2D ^{13}C - ^1H CPMG HSQC TOCSY, which provides excellent carbon resolution in the indirect dimension, as evident from Fig. S13 (ESI †).

We provide a complete set of experiments, which are CPMG and TOCSY based for both targeted and non-targeted analysis. The main aim was unambiguous identification and metabolic profiling. We combined the advantage of selective excitation for targeted analysis, multidimensional resolution for both proton and carbon, along with CPMG and TOCSY. We implemented TOCSY on HSQC, the most sensitive heteronuclear 2D experiment for metabolomics, along with CPMG to retain the benefit of enhanced carbon resolution. The experimental data



obtained using this technique leave all opportunities open for advanced pattern matching algorithms for assignment. To this end, Chenomx,³⁴ Amix, metaboanalyst,⁴¹ or COLMAR web server⁴² could easily be used individually or in combination depending upon the requirements. The peak list of both proton and carbon can narrow down the search criteria in the available standard databases, such as HMDB, BMRB, or MMCD.

Recent reports suggest that selective TOCSY is quantitative²⁶ and even more accurate because the selective experiment releases the spectral crowding.²⁶ Metabolic profiling commonly employs the quantification of any species based on the precise evaluation of signal integrals. The selective TOCSY experiment detects a single-spin system from the highly congested spectra of biofluid. As a consequence, the assessment of signal integral becomes more accurate. Selective TOCSY has previously been employed to obtain correct concentrations for a set of metabolites over a complete dynamic range in honey samples.²⁶ The absolute concentration determinations can be achieved by calibrating the TOCSY peak intensities using an external standard. A similar approach is used in 2D ¹H-¹H CPMG COSY also for BCAAs.² Based on these observations, we expect a positive quantitative ability for this minimal set of experiments.

Conflicts of interest

The authors declare that they have no conflict of interest.

Acknowledgements

The authors would like to thank the Imaging and Characterization Core Lab at the King Abdullah University of Science and Technology (KAUST) for access to the NMR facilities. This publication is based on work supported by KAUST Smart Health Initiative seeding grants (L. J. and M. J.) and through baseline-funds (L. J. and M. J.).

Notes and references

- 1 A. H. Emwas, R. Roy, R. T. McKay, L. Tenori, E. Saccenti, G. A. N. Gowda, D. Raftery, F. Alahmari, L. Jaremko, M. Jaremko and D. S. Wishart, *Metabolites*, 2019, **9**, 123.
- 2 S. Ghosh, A. Sengupta and K. Chandra, *Amino Acids*, 2015, **47**, 2229–2236.
- 3 P. S. Belton, I. J. Colquhoun, E. K. Kemsley, I. Delgadillo, P. Roma, M. J. Dennis, M. Sharman, E. Holmes, J. K. Nicholson and M. Spraul, *Food Chem.*, 1998, **61**, 207–213.
- 4 P. Sandusky and D. Raftery, *Anal. Chem.*, 2005, **77**, 2455–2463.
- 5 C. H. Johnson, J. Ivanisevic and G. Siuzdak, *Nat. Rev. Mol. Cell Biol.*, 2016, **17**, 451–459.
- 6 G. A. N. Gowda and D. Raftery, *Anal. Chem.*, 2014, **86**, 5433–5440.
- 7 H. Tang, Y. Wang, J. K. Nicholson and J. C. Lindon, *Anal. Biochem.*, 2004, **325**, 260–272.
- 8 P. G. Williams, J. K. Saunders, M. Dyne, C. E. Mountford and K. T. Holmes, *Magn. Reson. Med.*, 1988, **7**, 463–471.
- 9 L. Braunschweiler and R. R. Ernst, *J. Magn. Reson.*, 1983, **53**, 521–528.
- 10 M. Novakovic, Ě. Kupče, A. Oxenfarth, M. D. Battistel, D. I. Freedberg, H. Schwalbe and L. Frydman, *Nat. Commun.*, 2020, **11**, 5317.
- 11 R. Freeman and Ě. Kupče, *eMagRes*, DOI: 10.1002/9780470034590.emrstm1035.
- 12 P. Gierth, A. Codina, F. Schumann, H. Kovacs and Ě. Kupče, *Magn. Reson. Chem.*, 2015, **53**, 940–944.
- 13 O. Cloarec, M. E. Dumas, A. Craig, R. H. Barton, J. Trygg, J. Hudson, C. Blancher, D. Gauguier, J. C. Lindon, E. Holmes and J. Nicholson, *Anal. Chem.*, 2005, **77**, 1282–1289.
- 14 L. M. Smith, A. D. Maher, O. Cloarec, M. Rantalainen, H. Tang, P. Elliott, J. Stamler, J. C. Lindon, E. Holmes and J. K. Nicholson, *Anal. Chem.*, 2007, **79**, 5682–5689.
- 15 L. P. McIntosh, in *Encyclopedia of Biophysics*, ed. G. C. K. Roberts, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 386–386, DOI: 10.1007/978-3-642-16712-6_320.
- 16 S. P. Rucker and A. J. Shaka, *Mol. Phys.*, 1989, **68**, 509–517.
- 17 A. Bax and D. G. Davis, *J. Magn. Reson.*, 1985, **65**, 355–360.
- 18 A. J. Shaka, C. J. Lee and A. Pines, *J. Magn. Reson.*, 1988, **77**, 274–293.
- 19 M. H. Levitt, R. Freeman and T. Frenkiel, *J. Magn. Reson.*, 1982, **47**, 328–330.
- 20 D. W. Bearden, S. Macura and L. R. Brown, *J. Magn. Reson.*, 1988, **80**, 534–538.
- 21 C. Griesinger, G. Otting, K. Wuethrich and R. R. Ernst, *J. Am. Chem. Soc.*, 1988, **110**, 7870–7872.
- 22 J. S. Waugh, *J. Magn. Reson.*, 1982, **49**, 517–521.
- 23 S. J. Glaser and G. P. Drobny, in *Advances in Magnetic and Optical Resonance*, ed. W. S. Warren, Academic Press, 1990, vol. 14, pp. 35–58.
- 24 H. Kessler, H. Oschkinat, C. Griesinger and W. Bermel, *J. Magn. Reson.*, 1986, **70**, 106–133.
- 25 P. Sandusky and D. Raftery, *Anal. Chem.*, 2005, **77**, 7717–7723.
- 26 P. Sandusky, E. Appiah-Amponsah and D. Raftery, *J. Biomol. NMR*, 2011, **49**, 281–290.
- 27 C. Papaemmanouil, C. G. Tsiafoulis, D. Alivertis, O. Tzamaloukas, D. Miltiadou, A. G. Tzakos and I. P. Gerothanassis, *J. Agric. Food Chem.*, 2015, **63**, 5381–5387.
- 28 N. MacKinnon, P. T. While and J. G. Korvink, *J. Magn. Reson.*, 2016, **272**, 147–157.
- 29 H. Geen and R. Freeman, *J. Magn. Reson.*, 1991, **93**, 93–141.
- 30 P. Schanda, *Development and application of fast NMR methods for the study of protein structure and dynamics*, Université Joseph-Fourier, Grenoble I, 2007.
- 31 G. Wider, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1998, **32**, 193–275.
- 32 J.-M. Bohlen, M. Rey and G. Bodenhausen, *J. Magn. Reson.*, 1989, **84**, 191–197.
- 33 M. J. Thrippleton and J. Keeler, *Angew. Chem., Int. Ed.*, 2003, **42**, 3938–3941.



- 34 P. Mercier, M. J. Lewis, D. Chang, D. Baker and D. S. Wishart, *J. Biomol. NMR*, 2011, **49**, 307–323.
- 35 A. M. Weljie, J. Newton, P. Mercier, E. Carlson and C. M. Slupsky, *Anal. Chem.*, 2006, **78**, 4430–4442.
- 36 S. Ghosh, A. Sengupta and K. Chandra, *Anal. Bioanal. Chem.*, 2017, **409**, 6731–6738.
- 37 P. K. Mandal and A. Majumdar, *Concepts Magn. Reson., Part A*, 2004, **20**, 1–23.
- 38 M. Piotto, V. Saudek and V. Sklenár, *J. Biomol. NMR*, 1992, **2**, 661–665.
- 39 X. Zhang, M. Liu and X. Mao, in *Encyclopedia of Spectroscopy and Spectrometry*, ed. J. C. Lindon, G. E. Tranter and D. W. Koppenaal, Academic Press, Oxford, 3rd edn, 2017, pp. 128–136, DOI: 10.1016/B978-0-12-409547-2.10945-X.
- 40 F. Delaglio, G. S. Walker, K. A. Farley, R. Sharma, J. C. Hoch, L. W. Arbogast, R. G. Brinson and J. P. Marino, *Am. Pharm. Rev.*, 2017, **20**, 339681.
- 41 J. Xia and D. S. Wishart, *Curr. Protoc. Bioinf.*, 2016, **55**, 1.
- 42 K. Bingol, D.-W. Li, B. Zhang and R. Brüscheiler, *Anal. Chem.*, 2016, **88**, 12411–12418.

