Optimised collection of non-uniformly sampled 2D-HSQC NMR spectra for use in metabolic flux analysis

Jeeves, Mark; Roberts, Jennie; Ludwig, Christian

DOI:
10.1002/mrc.5089

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Jeeves, M, Roberts, J & Ludwig, C 2020, 'Optimised collection of non-uniformly sampled 2D-HSQC NMR spectra for use in metabolic flux analysis', *Magnetic Resonance in Chemistry*. https://doi.org/10.1002/mrc.5089

Received: 3 April 2020 | Revised: 6 August 2020 | Accepted: 18 August 2020

DOI: 10.1002/mrc.5089

WILEY

TUTORIAL

Optimised collection of non-uniformly sampled 2D-HSQC NMR spectra for use in metabolic flux analysis

Mark Jeeves[1] | **Jennie Roberts**[2] | **Christian Ludwig**[2]

[1]Henry Wellcome Building for Biomolecular NMR Spectroscopy, Institute of Cancer and Genomic Sciences, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

[2]Institute of Metabolism and Systems Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

Correspondence
Christian Ludwig, Institute of Metabolism and Systems Research, College of Medical and Dental Sciences, University of Birmingham. Birmingham, UK.
Email: c.ludwig@bham.ac.uk

Funding information
Wellcome Trust, Grant/Award Number: 208400/Z/17/Z; Institutional Strategic Support Award given to the; UHB Charitable Funds, Grant/Award Number: 17-3-846

Abstract

Nuclear magnetic resonance (NMR) spectroscopy is integral to metabolic studies; yet, it can suffer from the long acquisition times required to collect data of sufficient signal strength and resolution. The use of non-uniform sampling (NUS) allows faster collection of NMR spectra without loss of spectral integrity. When planning experimental methodologies to perform metabolic flux analysis (MFA) of cell metabolism, a variety of options are available for the acquisition of NUS NMR data. Before beginning data collection, decisions have to be made regarding selection of pulse sequence, number of transients and NUS specific parameters such as the sampling level and sampling schedule. Poor choices will impact data quality, which may have a negative effect on the subsequent analysis and biological interpretation. Herein, we describe factors that should be considered when setting up non-uniformly sampled 2D-^{1}H,^{13}C HSQC NMR experiments for MFA and provide a standard protocol for users to follow.

KEYWORDS

isotopomers, metabolism, MFA, NMR, NUS, tracer

1 | INTRODUCTION

Stable isotope tracing is routinely used to elucidate the activities of various metabolic pathways inside biological systems, from whole organisms to individual cells.[1–7] Metabolic substrates enriched with stable isotopes, such as ^{13}C labelled glucose and glutamine, are used to trace the metabolism of a system. Metabolic transformations distribute labelled atoms into downstream metabolites that can be detected using techniques such as mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. Combining data on overall incorporation of isotopes available from MS and NMR spectroscopy allows detailed examination of the fate of atoms throughout a metabolic system.[7]

NMR spectroscopy is uniquely able to give information on the position of isotopically labelled nuclei, in contrast to the more sensitive technique of MS, which can only provide the relative amounts of each isotope isomer (isotopomer). In ^{1}H decoupled 1D-^{13}C NMR spectra, the multiplet pattern of an NMR peak is a result of the constituent isotopomers (see Figure 1B). In order to use this information to calculate the corresponding contributions from each isotope position, additional information is needed. As 1D-^{13}C NMR spectroscopy is blind to ^{12}C, the multiplet information must be scaled to derive the relative contribution of unlabelled nuclei (i.e., molecules containing ^{13}C nuclei only at natural abundance). Such scaling can be provided either by determining the per carbon ^{13}C percentage[10–15] or by employing the

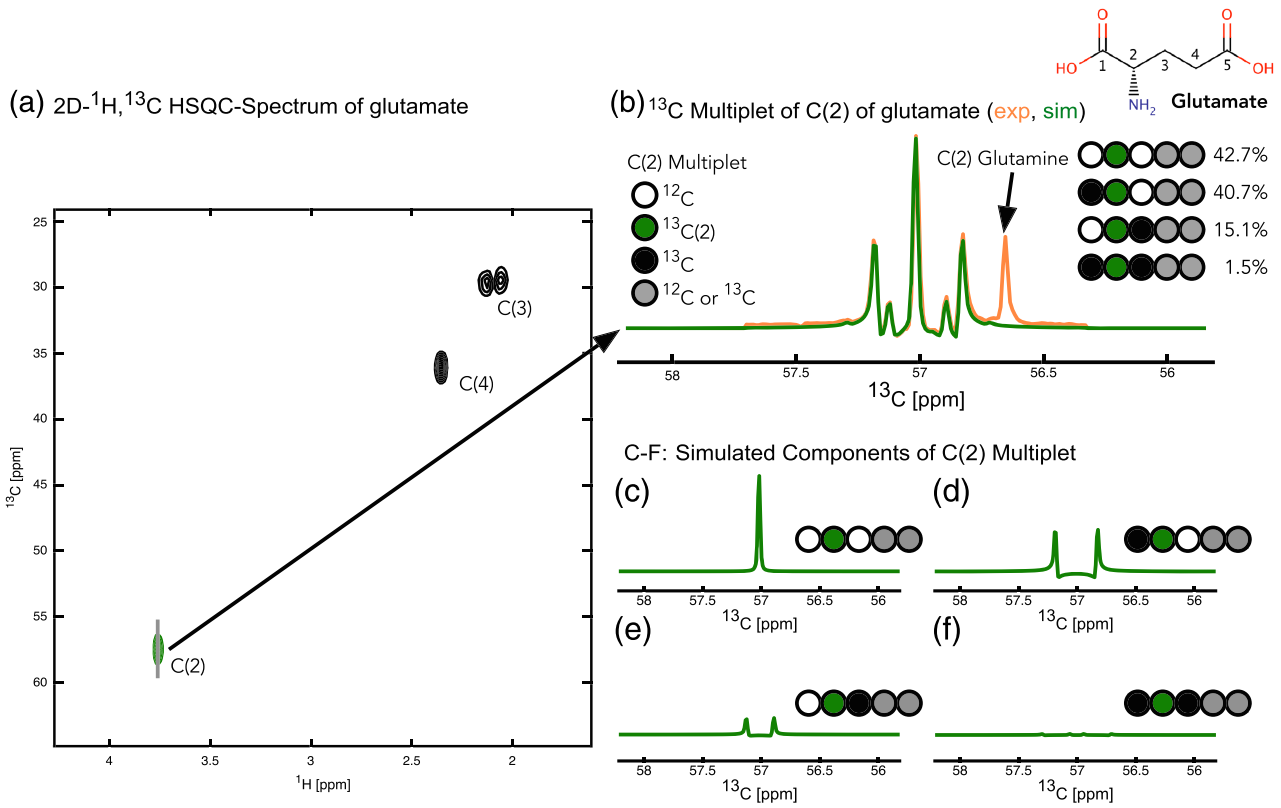

FIGURE 1 Schematic 2D-^{1}H,^{13}C HSQC spectrum of glutamate with multiplet analysis of the C(2) signal. Panel A shows the 2D-HSQC spectrum with the signals for C(2), due to chemical shift differences of the two protons attached to C(3) two signals for C(3) and the signal for C(4) all indicated. Panel B shows the experimental (amber) and simulated (green) multiplet obtained from the ^{13}C dimension of the C (2) signal with percentages of each potential isotopomer derived from fitting the intensity of the four possible multiplet components displayed in panels C–F to an experimental spectrum. The schematic representations of glutamate displayed next to each 1D spectrum show white circles for unlabelled carbon atoms, black circles for labelled carbon atoms and grey circles where carbon atoms can either be ^{12}C or ^{13}C without affecting signal multiplicity. C(2) of glutamate is always shown in green. The signal fitting was performed within the MetaboLab software,[8] using the pyGamma library[9] for NMR lineshape simulations. The 1D-^{13}C slice of the 2D-NMR spectrum was simulated as a 1D-^{13}C spin echo NMR spectrum with the echo time set to 1.55 ms to account for the gradient selection ^{13}C spin-echo within the 2D-^{1}H,^{13}C HSQC NMR pulse sequence

isotopomer distribution derived from MS data.[7] Once this scaling has been completed, the data from the multiplet patterns of the ^{13}C NMR spectra can be used to determine position specific ^{13}C distribution in each molecule.

The detected ^{13}C multiplet pattern can be interpreted directly, known as tracer analysis, or it can be integrated into a more demanding metabolic flux analysis (MFA) that requires metabolic network models to extract in vivo reaction rates.[3] For the remainder of this article, we will use the term MFA to include tracer analysis. MFA provides valuable biological insights, highlighting the relative activity of different metabolic pathways and the nutrient contribution to metabolites from labelled substrates.

The structure and labelling pattern of metabolites can be determined using a variety of NMR experiments. Positional enrichment of isotopes provides information on the active metabolic pathways, and this detailed isotopomer analysis can use a mixture of experiments. 1D-^{1}H spectra provide a wealth of information; however, the narrow frequency range results in significant resonance overlap that complicates peak integration. Isotopically enriched metabolites can be identified directly from the ^{13}C satellites in proton spectra,[14] or 1D-^{13}C NMR spectra can be used to elucidate positional isotopomers. However, the low sensitivity of ^{13}C nuclei and their longer relaxation times limits this approach to concentrated samples, even when ^{13}C-optimised probes are employed.[15]

Whereas 1D-^{1}H NMR experiments are often faster compared with 2D NMR experiments, the same is not necessarily true for 1D-^{13}C NMR experiments. Two-dimensional ^{1}H,^{13}C correlation NMR experiments are usually ^{1}H detected and are therefore more sensitive due to the different gyromagnetic ratios of ^{1}H and ^{13}C. In addition, 2D-^{1}H,^{13}C correlation NMR experiments reduce congestion of signals and aid identification by dispersion of

chemical shift into the two dimensions. The toolbox of NMR methods for isotopomer analysis includes acquisition techniques such as 2D-HSQC and HMBC NMR experiments, which correlate proton frequencies with the relevant heteronucleus (usually ^{13}C or ^{15}N in MFA experiments),[16] and the HSQC-TOCSY NMR experiment, which can be used to identify molecules with multiple labels.[17] Another possible approach used to speed up data acquisition is to reduce the relaxation delay between transients, and such fast 2D approaches have demonstrated that 2D-^{1}H,^{13}C correlation NMR spectra can be obtained using experiment times of less than half an hour.[18,19] Complete 2D correlations can be acquired in a single scan, and ultrafast zTOCSY and COSY techniques have been used to measure specific ^{13}C enrichments.[20]

Positional ^{13}C percentages can be derived from a variety of different NMR spectra such as various forms of 1D-^{1}H NMR spectra[10–12,14,15] or 2D-^{1}H,^{1}H TOCSY NMR spectra.[13] Spectral congestion can be a real problem for 1D-^{1}H NMR spectra. Although this can be resolved through the use of 2D-^{1}H,^{1}H TOCSY NMR spectra, there can still be substantial overlap, which can make it impossible to obtain fully quantitative data for some carbon atoms in key metabolites (e.g., C(2) of glutamate, glutamine and the glutamyl part in glutathione). The use of MS-derived mass isotopomer distributions in a full isotopomer analysis can overcome these problems.[7] In addition, the use of two orthogonal analytical technologies such as NMR spectroscopy and MS serves as an integrated consistency check. The use of a 2D-^{1}H,^{13}C HSQC NMR pulse sequence in combination with gas chromatography coupled MS (GC-MS)-based mass isotopomer distribution data is one of the most effective ways to trace the incorporation of ^{13}C nuclei into metabolites.[7] Using typical settings, the cross-peaks appearing in a 2D-^{1}H,^{13}C HSQC NMR spectrum are ^{1}H nuclei bonded directly to ^{13}C nuclei and yield well-resolved spectra. Figure 1 demonstrates the information content of a high-resolution 2D-^{1}H,^{13}C HSQC NMR spectrum. Whereas panel A depicts an overview plot of the 2D-^{1}H,^{13}C HSQC NMR spectrum of glutamate, panel B shows a high resolution slice through the ^{13}C multiplet (in amber) and a simulation of the multiplet line shape (green) of the indirect (^{13}C) dimension for the signal of C(2) of glutamate. The percentages indicate how much each component contributes to the overall multiplet. Panels C–F plot each multiplet component. A schematic representation of glutamate was added to each subpanels B–F with black circles indicating ^{13}C labelled nuclei, white circles representing ^{12}C nuclei. Because the signal of C(2) is shown here, C(2) is always ^{13}C and is shown in green. Circles shown in grey can be either ^{12}C or ^{13}C nuclei without affecting the multiplet of C(2) of glutamate.

As the length of the free induction decay (FID) in the ^{13}C dimension dictates the resolution of the ^{13}C dimension of the 2D NMR spectrum, it is necessary to acquire large numbers of increments resulting in a long experiment time. This is impractical when dealing with multiple samples; however, reductions in experimental times can be made with non-uniform sampling (NUS).[21] NUS schedules are designed to collect only a fraction of the data points after which reconstruction algorithms convert the NUS data into a full signal. NUS is therefore a vital technique in facilitating the acquisition of NMR data for metabolic tracing in a reasonable time frame. Further time savings can be achieved through enhancing the signal splitting due to J-coupling in the ^{13}C dimension, allowing a reduced number of increments to be acquired while still being able to resolve the individual components of the multiplets. Enhanced splitting also enables the use of higher field spectrometers, providing greater signal-to-noise ratio without the need to collect extra increments that would otherwise be required to resolve the splitting due to J-coupling.[15,22] NUS approaches have also been used to obtain high resolution NMR spectra without resolving ^{13}C–^{13}C coupling patterns to increase chemical shift accuracy and other coupling patterns in two-dimensional NMR spectra.[23,24]

There are a number of options to be considered when designing the experimental approach, from the choice of tracer to the selection of pulse sequences, sampling level or sampling schedule. Some of these options will have significant effects on data quality and precision. While we will not discuss options dictated by the biological problem, such as choice of tracer, we will explore NMR data acquisition-related options. Specifically, we will test the consequences of the choice of pulse sequence, number of transients acquired, sampling level and sampling schedule then explore the potential of using 2D-HSQC NMR spectra suitable for MFA studies in high-throughput settings (i.e., acquisition of 2D-^{1}H,^{13}C HSQC NMR spectra well enough resolved to analyse ^{13}C–^{13}C J-coupling patterns in under 15 min). We will use the results to recommend a standardised approach for the acquisition of non-uniformly sampled 2D-HSQC NMR data for MFA.

2 | PULSE SEQUENCES

Modern NMR spectrometers offer a variety of different pulse sequences for the acquisition of 2D-HSQC NMR spectra, each of which with its own advantages and disadvantages. We tested three different pulse sequences with respect to signal-to-noise ratio and artefact levels. The pulse sequences tested were a States-TPPI-based

sequence[25] and two echo/anti-echo (E/A)-based sequences,[26–28] one of which included sensitivity enhancement, and the other pulse sequence did not. Initially, 2D-^{1}H,^{13}C HSQC NMR spectra were acquired with full sampling, to separate the influence of pulse sequence on the measured parameters from artefacts generated by the use of NUS. In order to limit experimental time, two transients were used for 8,192 complex increments.

While the spectra acquired using an E/A-based pulse sequence are virtually artefact free (Figure 2A), the use of States-TPPI for quadrature detection results in significant artefacts especially in the crowded regions of the spectrum (Figure 2B). In addition to artefacts that look similar to t_1 noise, artefactual signals can be observed in the middle and at the edges of the spectrum due to insufficient phase cycling. As a consequence, it is impossible to faithfully detect signals from nuclei such as the C(2) of alanine, even though the presence of splitting in the C (3) of alanine shows that it must be labelled and therefore should be observable in the NMR spectrum. Another consequence of the insufficient artefact suppression using a States-TPPI-based pulse sequence with only two transients is that it is impossible to phase all the NMR signals in the spectrum simultaneously. For example, the methyl group signals of alanine and lactate are about 7.2° out of phase in the ^{13}C dimension even though they are in

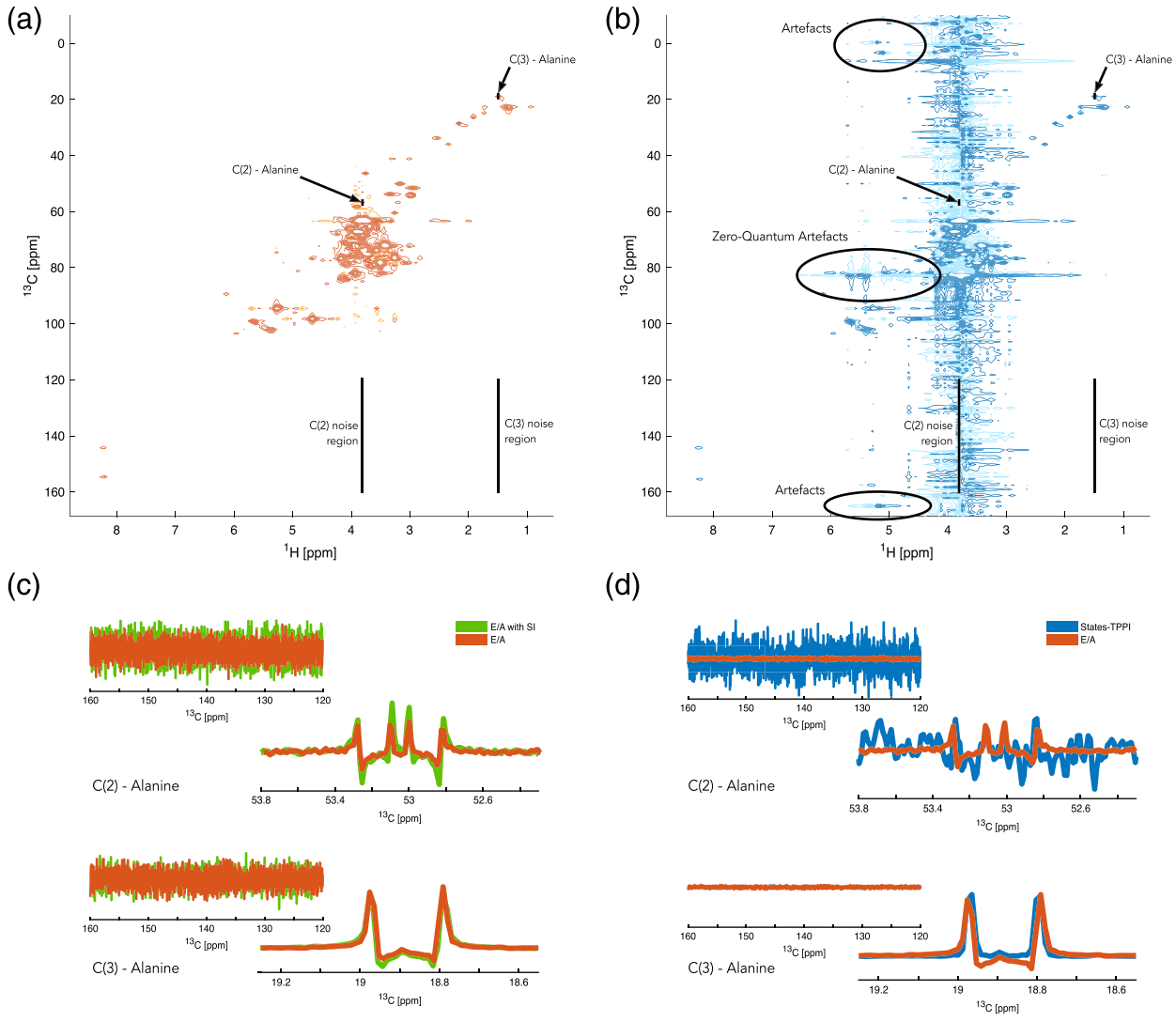

FIGURE 2 Fully sampled 2D-^{1}H,^{13}C HSQC NMR spectra of a sample of cell extract collected using either E/A (A) or States-TPPI (B) for quadrature detection. The spectra were acquired without enhancement of the ^{13}C–^{13}C splittings. Peaks from the detectable carbons of alanine are indicated along with regions selected for noise calculation. Artefacts that are not edited out in the phase cycle are highlighted. The ^{13}C traces for the C (2) and C (3) carbons of alanine are shown for spectra collected using E/A both with and without sensitivity improvement together with the traces of regions without peaks to allow estimation of noise levels (C). The overlay of ^{13}C traces for the C (2) and C (3) carbons of alanine collected with States-TPPI and E/A are shown together with the noise region (D). Overlaid spectra are shown with the same peak intensity relative to each other

spectral proximity. Because the analysis of 2D-HSQC NMR spectra in MFA studies critically depends on proper phasing of the spectrum, this will prevent a quantitative analysis of such spectra. The use of E/A quadrature detection results in much improved artefact suppression even when using just two transients. The slightly twisted line shape, due to the homonuclear ^{13}C–^{13}C J-coupling contribution to the first increment in the ^{13}C dimension, must not be confused with the differences in phase observed for several signals in the States-TPPI spectrum. E/A HSQC spectra can be easily phased, and as the signal distortion is predictable, it can be accounted for in a quantitative analysis such as a line shape simulation. This makes E/A-based pulse sequences ideal for collection of 2D-^{1}H,^{13}C HSQCs NMR spectra with high resolution in the ^{13}C dimension.

The signal from areas of the spectrum where no resonances are expected can be used as a measure of both the noise and the artefact levels present. It can clearly be seen that the use of E/A for quadrature detection results in significantly smaller signals in the empty regions of the spectrum confirming the smaller numbers of artefacts compared with the States-TPPI (Figure 2).

Sensitivity enhancement used in conjunction with E/A results in the expected modifications in peak intensity with an increase in peak intensity for CH groups but no enhancement for CH_3 (Figure 2C). However, there is an increase in the J-coupling contribution in the first increment, which results in a change in line shape. In protein NMR spectroscopy, the observed groups often have a common quality, for instance the amide group of the peptide bond, where the number of protons attached to nitrogen is one. In contrast, small molecules (e.g., metabolites) have a range of different groups, and the numbers of protons attached to the heavy atoms vary, so to cover the widest range of observable signals, the use of sensitivity enhancement is not recommended.

In summary, E/A is the most suitable variant of 2D-HSQC NMR pulse sequences for measuring 2D-HSQC NMR spectra with high resolution in the ^{13}C dimension due to the quality of data acquired with a phase cycle that requires only two transients.. The lack of artefacts present in the spectra acquired using E/A lends itself to being used in conjunction with NUS as artefacts present in the normally acquired spectra will be amplified by the use of NUS. A quantitative analysis of such spectra necessitates the use of line shape simulations to account for the presence of ^{13}C–^{13}C J-coupling contributions to the first increment in the ^{13}C dimension. The use of States-TPPI inevitably leads to increased artefact levels, which may bias data analysis of such spectra. In addition, we observed that some areas of the State-TPPI spectrum are unusable due to the insufficient filtering of residues from large peaks (C(2) of alanine, for example).

3 | INFLUENCE OF NUMBER OF TRANSIENTS ON SIGNAL-TO-NOISE RATIO

When collecting NMR data, there is always a balance between the time taken to acquire the spectra and the signal-to-noise ratio that can be achieved. Metabolic tracing samples are often mass limited; therefore, various measures must be taken to optimise signal-to-noise. Probes designed to handle very low volumes are used to make best use of the sample available together with the greater sensitivity of the highest field magnets available. The 1.7-mm TCI CryoProbe working at 800 MHz is the best set-up currently available. While increasing the number of transients will lead to stronger signals, the time taken may be prohibitively long and result in a reduction of the number of samples that can be run in any given period. The E/A 2D-HSQC NMR pulse sequence can be used with a minimum of two transients. However, does this result in enough signal, or could increasing the number of transients give better data and also result in a decrease in artefacts? As demonstrated in Figure 3, there is no appreciable increase in artefacts when using two transients as opposed to four or eight (Figure 3A,C). All the fine features of the peak multiplets are present, and there is only a small increase in the signal-to-noise seen when increasing from two transients to eight. The benefit from collecting more transients when using NUS does not result in the same increase in signal-to-noise as with fully sampled spectra, and a better use of time may be to increase the sampling level. The case with the States-TPPI 2D-HSQC NMR pulse sequence is different. While isolated signals, such as the methyl group of alanine (Figure 3D), follow the same pattern as signals in the E/A 2D-HSQC NMR spectrum (Figure 3A, C), signals in areas of the spectrum with signals from highly labelled molecules such as glucose, which possess signals with a distinctive ^{13}C but very similar ^{1}H chemical shifts, may not be observable due to spectral artefacts overshadowing these signals. An example is demonstrated in Figure 3B, where the equivalent trace of the two transient E/A 2D-HSQC NMR spectrum was plotted as well.

It is important to note that the main issue arising from the choice of pulse sequences is not the signal-to-noise but rather the signal-to-artefact ratio. The presence of artefacts (which can appear as noise) severely limits the use of the data that can compromise the resultant analysis. In summary, it confirms that even if the full

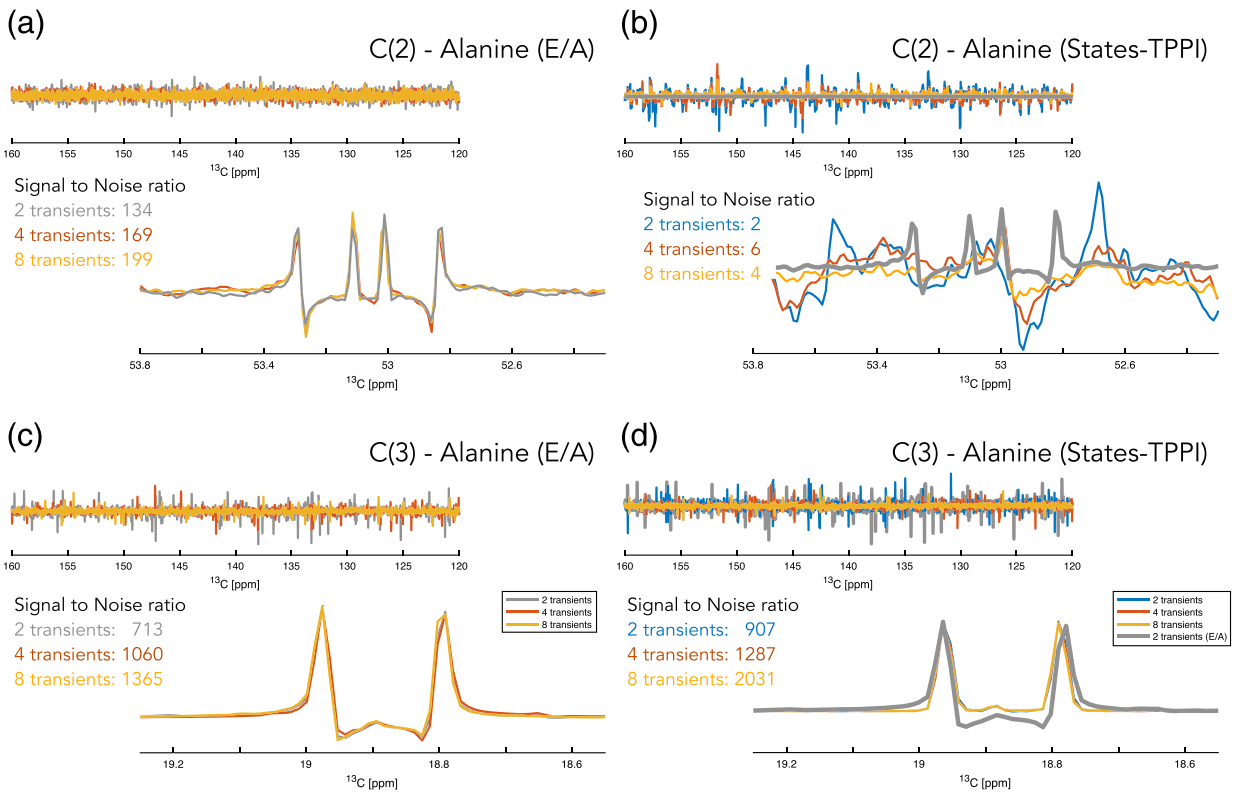

FIGURE 3 The effect of the number of transients on the signals acquired using E/A and States-TPPI pulse sequences. ^{13}C slices from the 25% non-uniformly sampled 2D-^{1}H,^{13}C HSQC NMR spectra showing the alanine multiplets corresponding to C(2) (A,B) and C(3) (C,D). The region from the same slice where no peaks are expected is used to illustrate noise levels. The traces for spectra acquired with 2, 4 and 8 transients are shown. The traces from the States-TPPI spectra are overlaid with same trace from the E/A spectrum collected with 2 transients (grey) in order to show the position of missing peaks

phase cycle is used, the States-TPPI 2D-HSQC NMR pulse sequence is not suitable for MFA studies and that the E/A-based 2D-HSQC NMR pulse sequences are the most suitable for this purpose because of the efficient filtering of artefacts from the collected spectra and can be used with only two transients per increment.

It should be noted that all non-uniformly sampled spectra presented in this study were reconstructed using the IRLS[29,30] algorithm with 20 iterations using the NMRPipe[31] and qMDD[30,32] software packages. There was no difference found in spectra reconstructed using iterative soft thresholding (IST)[33,34] with 400 iterations and those using IRLS so either approach is applicable to MFA studies.

4 | MINIMUM SAMPLING LEVEL REQUIRED

As we routinely collect 8,192 complex increments in the fully sampled spectra, even a relatively low sampling level will result in a relatively large number of points being collected. As discussed above, the sampling level may be kept high to maximise signal-to-noise but if the amount of sample is high enough such that signal is not an issue then what is the minimum sampling level required? There is no noticeable increase in artefacts when reducing the sampling from 50% to 20% and the ratio of signal-to-the-noise region of the spectrum remains stable (Figure 4). However, when the sampling level is dropped to below 20%, some fine features of the multiplet are lost, and so, the use of very low sampling levels is not recommended for general use if the dynamic range of observable signals is not to be limited. The isotope distributions calculated from data with various sampling levels (Table 1) are similar. However, the minor contributions of the $^{13}C(1)$,$^{13}C(2)$,$^{12}C(3)$ and $^{12}C(1)$,$^{13}C(2)$,$^{13}C(3)$ components of the C(2) lactate multiplet are lost at sampling levels of 10% or lower and are underestimated at 20%, which suggests that for most experiments a sampling level of below 25% is not recommended.

In summary, for 2D-HSQC NMR spectra, an NUS schedule that acquires 25% of the 8,192 complex data points of the ^{13}C dimension gives the best solution allowing rapid collection of data without the loss of spectral integrity.

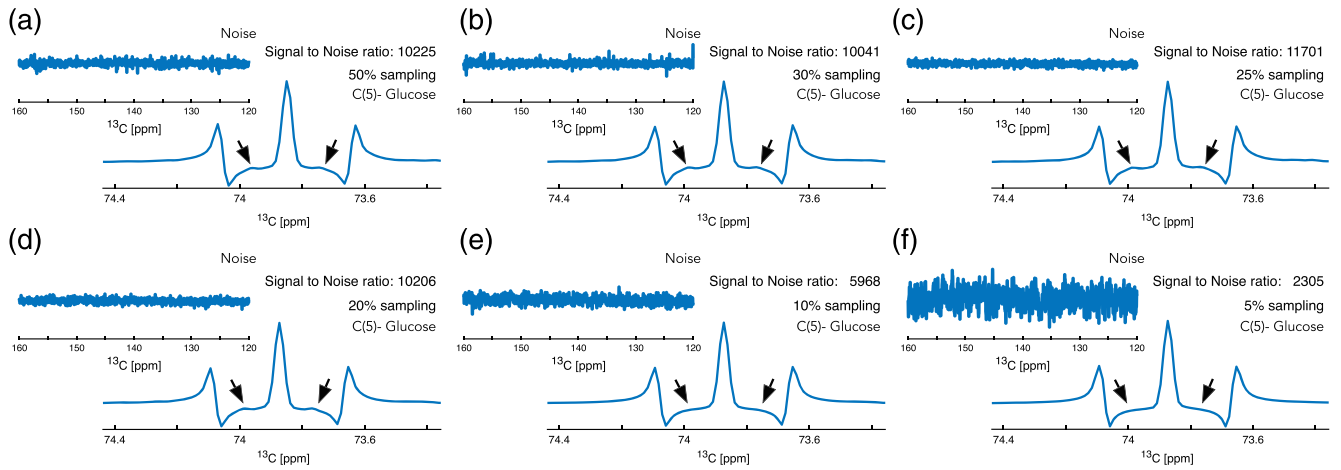

FIGURE 4 The effect of varying the sampling level on the multiplet signal of C(5) of glucose. The ^{13}C trace of the 2D-^{1}H,^{13}C HSQC NMR spectra collected using sampling levels of 50% (A), 30% (B), 25% (C), 20% (D), 10% (E) and 5% (F) are show together with the corresponding noise from a range where no signals are expected. Arrows indicate fine features of the multiplet that are lost with lower sampling levels

TABLE 1 Contributions of each component to the overall multiplets of lactate with different sampling levels as measured by simulation of the multiplet

Lactate C(2) observed				
Sampling	**$^{12}C(1)$, $^{13}C(2)$, $^{12}C(3)$**	**$^{13}C(1)$, $^{13}C(2)$, $^{12}C(3)$**	**$^{12}C(1)$, $^{13}C(2)$, $^{13}C(3)$**	**$^{13}C(1)$, $^{13}C(2)$, $^{13}C(3)$**
100%	2	3	3	92
50%	2	3	3	92
30%	2	3	3	92
25%	2	3	3	92
20%	2	2	2	94
10%	2	0	0	98
5%	2	0	0	98
Lactate C(3) observed				
Sampling	**$^{12}C(2)$, $^{13}C(3)$**		**$^{13}C(2)$, $^{13}C(3)$**	
100%	2		98	
50%	2		98	
30%	2		98	
25%	2		98	
20%	2		98	
10%	2		98	
5%	2		98	

Note: 2D-^{1}H,^{13}C HSQC NMR spectra with NUS levels of 100%, 50%, 30%, 25%, 20%, 10% and 5% were acquired, and the contributions of each component of the multiplet were calculated based on the comparison of the real data with simulations of the multiplet with varying component contributions. Each component relates to a different isotope incorporation into the carbons of lactate. The line-shape simulations were performed within the MetaboLab software[20] using the pyGamma library.[9] C(1) of lactate is not observable by 2D-^{1}H,^{13}C HSQC NMR spectroscopy as it has no attached proton. No information on the isotopic state of C(1) can be assessed from observation of C(3) as the two bond ^{13}C–^{13}C J-coupling is too small to be resolved.

5 | SAMPLING SCHEDULES

Sampling schedules can make a large difference to data for protein samples where weighting the distribution of the increments sampled based on the transverse relaxation time (T_2) of the sample can have dramatic effects.[35,36] With samples of small molecules, for example, most polar metabolites, the T_2 of these

molecules is quite long, and as such, varying the sampling schedule should not influence the resulting spectrum as much.[37] Indeed, we did not find any noticeable difference between NMR spectra acquired using either sampling schedules weighted towards the initial increments (Figure 5A,B) versus non-weighted (Figure 5C) or a sampling schedule weighted towards the beginning and the end of the indirect evolution period (Figure 5D). As there is no effect of using different sampling schedules, we do not recommend any particular sampling scheme; they should all lead to equivalent results. All sampling schedules, except those used to produce the data for Figure 5, were weighted towards the initial increments.

6 | SUITABILITY FOR HIGH-THROUGHPUT MFA STUDIES

Observation of metabolism in real time is becoming increasingly popular.[38–43] Usually, 1D NMR pulse sequences or hyperpolarized samples are used to acquire metabolic data from living samples. Although hyperpolarisation is a truly fascinating technique, it requires highly specialised equipment, beyond the capabilities of most labs. Proton-detected 1D NMR pulse sequences can provide metabolic tracing information on a rapid timescale, however quantitative data can only be obtained for sparse NMR spectra as spectral congestion results in peak overlap.[14,15,42,43] We

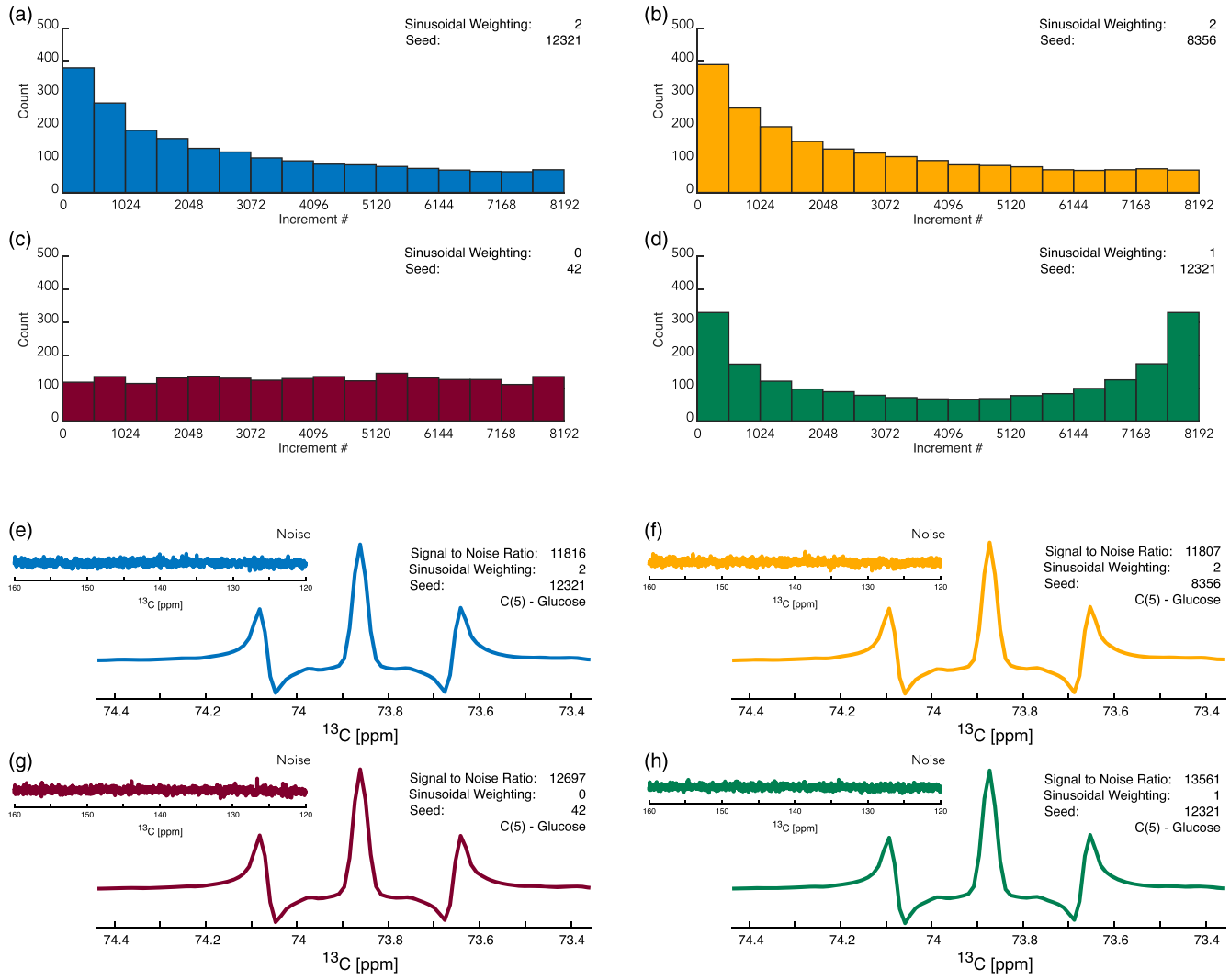

FIGURE 5 The effect of weighting the sampling of 25% non-uniformly sampled 2D-^{1}H,^{13}C HSQC NMR spectra. Two front-end weighted sampling schedules starting from different seeds were tested (A,B) with the ^{13}C slice of the 2D-HSQC NMR spectrum corresponding to C(5) of glucose is shown together with a corresponding noise region where no signals are expected (E,F). Non-weighted (C,G) and front and back-end weighted (D,H) 25% non-uniformly sampled 2D-HSQC NMR spectra were also acquired

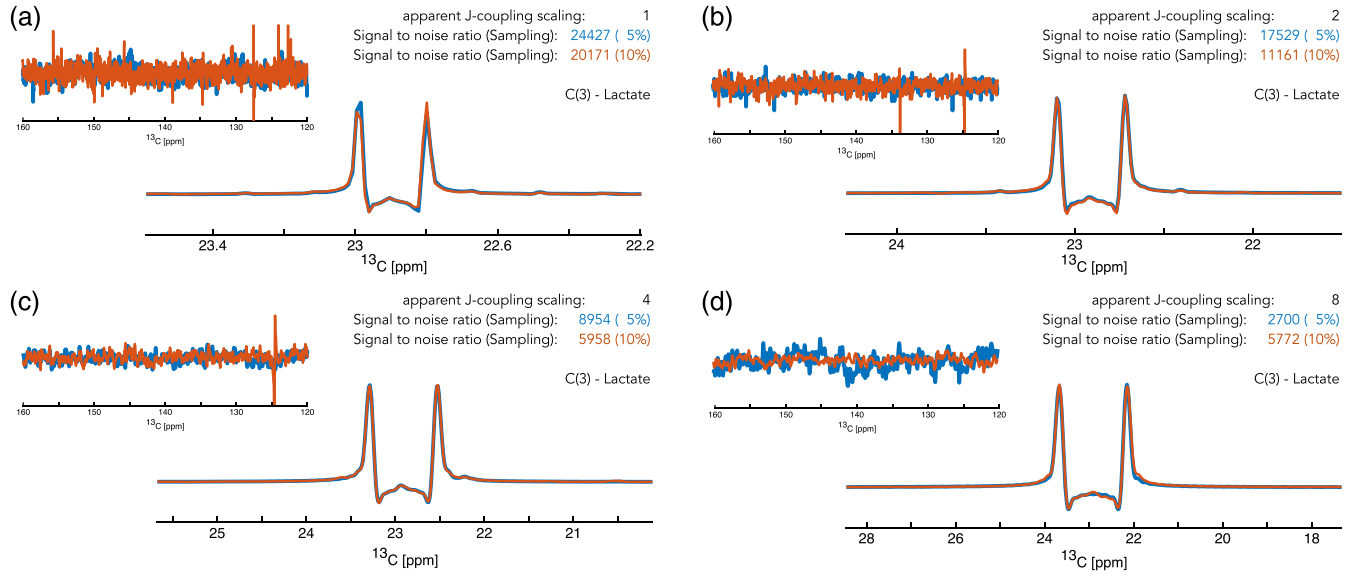

FIGURE 6 The utility of non-uniformly sampled high resolution 2D-HSQC NMR spectra to high-throughput NMR studies. ^{13}C-traces of 5% non-uniformly sampled 2D-^{1}H,^{13}C HSQC NMR spectra collected with J-coupling splitting enhancement of 1 (A), 2 (B), 4 (C) and 8 (D) together with a corresponding noise region

therefore explored the possibility of acquiring 2D-HSQC NMR spectra at rates suitable for high-throughput data acquisition.

Using small NUS levels (5% or 10%), in combination with the enhancement of apparent signal splittings,[15,22] it is possible to acquire 2D-HSQC NMR spectra with enough spectral resolution in the ^{13}C dimension to resolve ^{13}C–^{13}C J-coupling-based signal splitting in under 15 min (Figure 6). This is only possible for a limited number of metabolites with large and well-resolved signals, such as lactate, due the limitations of the observable dynamic range. This means that the large signals are reconstructed successfully, but the lack of data points results in poor reconstruction of smaller signals, often resulting in them being absent in the reconstructed spectrum. Increasing the J-coupling splitting enhancement to eight resulted in the loss of some features as seen by the disappearance of central peak of the multiplet (Figure 6D). As there is little else to differentiate between a spectrum collected with an enhancement of 4% and 5% sampling level and a spectrum collected with an enhancement of 8% and 10% sampling level (Table 2), and as both spectra take the same amount of time to acquire, it is better to limit the enhancement of the splitting due to J-coupling to 4% and instead reduce the sampling level to 5% in order to retain all features of the multiplets. Rapid collection of NMR data suitable for MFA studies is therefore possible, with appropriate attention to the parameters used.

7 | SUMMARY

This article describes some of the considerations that need to be taken into account before setting up a tracer-based experiment using 2D-HSQC NMR spectroscopy. The limitation of the low sensitivity of NMR spectroscopy compared with MS is counteracted by the richness of the data that can be obtained. There is a need to speed up NMR data acquisition, while still acquiring high-quality data of sufficient signal strength, in order to make NMR a viable tool for MFA studies. NUS plays a vital role in this, but it must be used with the correct parameters in order to yield meaningful data.

The choice of pulse sequence is vital in the elimination of artefacts that compromise data analysis when using NUS. Collecting data with very few repetitions is most effectively achieved with the use of echo/anti-echo for coherence selection, as only two transients are required for artefact removal.

While striving for data acquisition in the shortest possible time, it is important not to reduce the number of points sampled such that the quality of the data is compromised. Equally, there is no need to sample more points than necessary, as this achieves little and wastes time. The use of 25% sampling provides the ideal compromise between rapid data acquisition and high-quality data output. The weighting of the non-uniformly sampled points has no impact on spectra for MFA due to the small relaxation rates in small molecules. In the emerging field of high-throughput metabolic analysis, it is possible, with

a very low NUS level, to acquire data rapidly enough to make NMR spectroscopy a viable tool in this field.

This article focuses on multiplet analysis, but it is worth noting that full isotopomer analysis requires the use of MS data for accurate results.[7] MS can monitor all isotopomers, unlike NMR data, which may lack certain isotopomer information. For example, with a standard 2D-^{1}H,^{13}C HSQC approach, quaternary carbons do not yield an NMR signal. In addition, NMR is blind to ^{12}C; therefore, it is not possible to directly quantify the amount of ^{13}C relative to ^{12}C at any specific position using a single sample. The combination of NMR multiplet data with MS isotopomer data resolves this problem.[7]

In summary, NUS is essential for the rapid acquisition of 2D-HSQC NMR spectra, allowing NMR spectroscopy to be a viable method for studying metabolism. The implementation of NUS, together with the low transient requirements of E/A and other methods such as enhancement of splitting due to J-coupling,[15,22] allows the acquisition of NMR data in a timely fashion.

8 | STANDARD PROCEDURE

Cells are cultured in the presence of isotopically labelled tracer, with the choice of tracer dependent on what particular pathways are of interest. In the absence of a priori knowledge, a metabolomics study can give insights into which pathways may be of interest to a particular system.

Metabolites are extracted using a biphasic system to separate polar molecules from lipids and remove other interfering molecules such as proteins.[44] The sample is split into portions for NMR (90%) and GC-MS (10%).

GC-MS is used to generate data giving number of incorporated heavy atoms in each metabolite, known as mass isotopomer distributions.[45,46] These data are not adequately provided by NMR alone as such a strategy would require the use of two samples in order to determine absolute fractional ^{13}C enrichment and thus lead to errors due to difficulties in reproducibility.

Two-dimensional HSQC NMR spectra are collected, and multiplet analysis is performed to determine position specific isotope incorporation, which cannot be achieved by MS. The standard set-up for the NMR spectroscopy suggested here involves collection of 2D-HSQC NMR spectra with the use of E/A for coherence selection in order to minimise artefacts. Full information on the experimental set-up used in this study is given in the methods section. At higher magnetic field stregnths, it will become necessary to enhance the splitting due to J-coupling in order to completely resolve the multiplet.[15,22] A Bruker pulse sequence capable of scaling the appearance of J-coupling-based signal splitting can be found in the data repository associated with this publication (http://doi.org/10.17605/OSF.IO/QTMGE).

The isotopologue data from GC-MS and the atom-specific incorporation data from NMR spectroscopy are combined for full isotopomer analysis.[7]

9 | METHODS

All 2D-HSQC NMR spectra were acquired using a Bruker NEO 800-MHz NMR spectrometer equipped with a 1.7-mm z-PFG TCI CryoProbe. The 2D-HSQC NMR spectra with echo/anti-echo gradient coherence selection were acquired using our previously described pulse sequence.[15,22] A presaturation pulse is used to suppress the water resonance during the 1.5-s interscan relaxation delay. The ^{1}H dimension in all experiments was acquired with a spectral width of 15.6 ppm using 512 complex data points. In the fully sampled spectra (Figure 2), the ^{13}C dimension was acquired with a spectral width of 189.8 ppm using 8,192 complex data points resulting in an experiment time of approximately 16 h. In spectra collected using NUS without splitting enhancement (Figures 3 and 5, Table 1), 25% of 8,192 complex data points were collected in the ^{13}C dimension resulting in an experiment time of approximately 4 h. In spectra to test the effect of NUS sampling level without the use of signal splitting enhancement (Figure 4), either 5%, 10%, 20%, 25% or 50% of 8,192 complex data points were collected in the ^{13}C dimension resulting in experiment times of approximately 0.8, 1.6, 3.2, 4 and 8 h, respectively. For experiments acquired using scaling of the signal splittings due to J-coupling in addition to NUS (Figure 6, Table 2), the following set-up was used: Experiments with no scaling were collected using either 5% or 10% of 8,192 complex data points, experiments using twofold scaling were collected using either 5% or 10% of 4,096 complex data points, experiments using fourfold scaling were collected using either 5% or 10% of 2,048 complex data points and experiments using eightfold scaling were collected using either 5% or 10% of 1,024 complex data points in the ^{13}C dimension[15,22] resulting in experiment times of 52, 104, 26, 52, 13, 26, 6.5 and 13 min, respectively. Fully sampled spectra were processed using NMRPipe.[31] The non-uniformly sampled spectra were reconstructed using the IRLS algorithm with 20 iterations within MDDNMR[30,32] (version 2.5) and processed using NMRPipe[31] (version 9.2). All spectra were processed with polynomial baseline correction after manual phase correction.

DATA AVAILABILITY STATEMENT

All experimental data for this article is available at http://doi.org/10.17605/OSF.IO/QTMGE. Experimental

TABLE 2 The effect of increasing the enhancement of the splitting due to J-coupling on the contributions of each component to the multiplets of lactate C(2) and C(3) as measured by simulation of the multiplet

Lactate C(2) observed					
J-coupling splitting enhancement	**NUS sampling level**	**^{12}C(1), ^{13}C(2), ^{12}C(3)**	**^{13}C(1), ^{13}C(2), ^{12}C(3)**	**^{12}C(1), ^{13}C(2), ^{13}C(3)**	**^{13}C(1), ^{13}C(2), ^{13}C(3)**
1	5%	2	0	0	98
2	5%	3	0	0	97
4	5%	2	0	0	98
8	5%	1	0	0	99
1	10%	2	0	0	98
2	10%	1	0	0	99
4	10%	2	0	0	98
8	10%	1	0	0	99
Lactate C(3) observed					
J-coupling splitting enhancement	**NUS sampling level**		**^{12}C(2), ^{13}C(3)**	**^{13}C(2), ^{12}C(3)**	
1	5%		2	98	
2	5%		2	98	
4	5%		2	98	
8	5%		0	100	
1	10%		2	98	
2	10%		1	99	
4	10%		2	98	
8	10%		1	99	

Note: 2D-^{1}H,^{13}C HSQC NMR spectra with NUS levels of 5% and 10% were collected with either 1-, 2-, 4- or 8-times enhancement of splitting due to J-coupling implemented. The contributions of each component to the multiplet were calculated by comparing the simulated multiplet with the real data. The line-shape simulations were performed within the MetaboLab software[8] using the pyGamma library.[9] C(1) of lactate is not observable by 2D-^{1}H,^{13}C HSQC NMR spectroscopy as it has no attached proton. No details of the status of C(1) lactate can be determined by the observation of C(3) as the splitting due to J-coupling is too small to be resolved.

datasets are as follows: pulseSequences_fullSampling.zip contains the fully sampled NMR spectra with E/A (subdirectory 1), States-TPPI (subdirectory 2) and E/A with sensitivity enhancement (subdirectory 3). The file transients_sampling25.zip contains spectra with 25% NUS with 2, 4 and 8 transients, respectively. Subdirectories 1–3 contain E/A 2D-HSQC NMR spectra, and subdirectories 4–6 contain States-TPPI 2D-HSQC NMR spectra. samplingRates_transients2.zip contains spectra exploring how much sampling is needed using an E/A 2D-HSQC NMR pulse sequence. Subdirectories 1–6 contain NMR spectra with sampling rates of 50%, 30%, 25%, 20%, 10% and 5%, respectively. The file samplingSchedules_sampling25_transients2.zip contains NMR spectra with different sampling schedules, and the file realTimeCapabilities.zip contains NMR spectra to explore the possibility of real-time acquisition of 2D-HSQC NMR spectra with sufficient resolution to be able to observe ^{13}C–^{13}C J-couplings. Please see the file metainformation.txt in the repository for details. The file pulseSequence.zip contains an E/A 2D-HSQC NMR pulse sequence with the possibility to scale the appearance of ^{13}C–^{13}C J-coupling constants in the indirect dimension. The MetaboLab software package is available at https://www.ludwiglab.org/software-development.

ACKNOWLEDGEMENTS

This work was supported in part by the Wellcome Trust (grant number 208400/Z/17/Z) and through an Institutional Strategic Support Award given to the University of Birmingham, UHB Charitable Funds (17-3-846). We thank the University of Birmingham for providing open access to their Wellcome Trust-funded 800-MHz spectrometer, the Metabolic Tracer Analysis Core (MTAC) at the University of Birmingham and Thomas Brendan Smith for preparing the sample used in this study.

ORCID

Christian Ludwig https://orcid.org/0000-0001-8901-6970

REFERENCES

[1] K. Hiller, C. M. Metallo, *Curr. Opin. Biotechnol.* **2013**, *24*(1), 60. https://doi.org/10.1016/j.copbio.2012.11.001

[2] J. Adam, M. Yang, C. Bauerschmidt, M. Kitagawa, L. O'Flaherty, P. Maheswaran, G. Özkan, N. Sahgal, D. Baban, K. Kato, K. Saito, K. Iino, K. Igarashi, M. Stratford, C. Pugh, D. A. Tennant, C. Ludwig, B. Davies, P. J. Ratcliffe, M. El-Bahrawy, H. Ashrafian, T. Soga, P. J. Pollard, *Cell Rep.* **2013**, *3*(5), 1440. https://doi.org/10.1016/j.celrep.2013.04.006

[3] J. M. Buescher, M. R. Antoniewicz, L. G. Boros, S. C. Burgess, H. Brunengraber, C. B. Clish, R. J. DeBerardinis, O. Feron, C. Frezza, B. Ghesquiere, E. Gottlieb, K. Hiller, R. G. Jones, J. J. Kamphorst, R. G. Kibbey, A. C. Kimmelman, J. W. Locasale, S. Y. Lunt, O. D. Maddocks, C. Malloy, C. M. Metallo, E. J. Meuillet, J. Munger, K. Nöh, J. D. Rabinowitz, M. Ralser, U. Sauer, G. Stephanopoulos, J. St-Pierre, D. A. Tennant, C. Wittmann, M. G. Vander Heiden, A. Vazquez, K. Vousden, J. D. Young, N. Zamboni, S.-M. Fendt, *Curr. Opin. Biotechnol* **2015**, *34*, 189. https://doi.org/10.1016/j.copbio.2015.02.003

[4] C. Lussey-Lepoutre, K. E. R. Hollinshead, C. Ludwig, M. Menara, A. Morin, L.-J. Castro-Vega, S. J. Parker, M. Janin, C. Martinelli, C. Ottolenghi, C. Metallo, A.-P. Gimenez-Roqueplo, J. Favier, D. A. Tennant, *Nat. Commun.* **2015**, *6*, 8784. https://doi.org/10.1038/ncomms9784

[5] G. M. Mackay, L. Zheng, N. J. F. van den Broek, E. Gottlieb, *In Methods in Enzymology; Elsevier* **2015**, *561*, 171. https://doi.org/10.1016/bs.mie.2015.05.016

[6] J. Nath, T. Smith, A. Hollis, S. Ebbs, S. Canbilen, D. Tennant, A. Ready, C. Ludwig, *Transpl. Res* **2016**, *5*(7). https://doi.org/10.1186/s13737-016-0037-0

[7] M. Chong, A. Jayaraman, S. Marin, V. Selivanov, P. R. de Atauri Carulla, D. A. Tennant, M. Cascante, U. L. Günther, C. Ludwig, *Angew. Chem., Int. Ed.* **2017**, *56*(15), 4140. https://doi.org/10.1002/anie.201611634

[8] C. Ludwig, U. L. Gunther, *BMC Bioinformatics* **2011**, *12*(1), 366. https://doi.org/10.1186/1471-2105-12-366

[9] S. A. Smith, T. O. Levante, B. H. Meier, R. R. Ernst, *J. Magn. Reson. A* **1994**, *106*(1), 75. https://doi.org/10.1006/jmra.1994.1008

[10] A. A. de Graaf, M. Mahle, M. Möllney, W. Wiechert, P. Stahmann, H. Sahm, *J. Biotechnol.* **2000**, *77* (1), 25. https://doi.org/10.1016/S0168-1656(99)00205-9

[11] A. P. Alonso, H. Vigeolas, P. Raymond, D. Rolin, M. Dieuaide-Noubhani, *Plant Physiol.* **2005**, *138*(4), 2220. https://doi.org/10.1104/pp.105.062083

[12] W. Eisenreich, J. Slaghuis, R. Laupitz, J. Bussemer, J. Stritzker, C. Schwarz, R. Schwarz, T. Dandekar, W. Goebel, A. Bacher, *Proc. Natl. Acad. Sci.* **2006**, *103*(7), 2040. https://doi.org/10.1073/pnas.0507580103

[13] A. N. Lane, T. W.-M. Fan, M. Bousamra, R. M. Higashi, J. Yan, D. M. Miller, *Omics J. Integr. Biol.* **2011**, *15*(3), 173. https://doi.org/10.1089/omi.2010.0088

[14] M. Vinaixa, M. A. Rodríguez, S. Aivio, J. Capellades, J. Gómez, N. Canyellas, T. H. Stracker, O. Yanes, *Angew. Chem., Int. Ed.* **2017**, *56*(13), 3531. https://doi.org/10.1002/anie.201611347

[15] T. B. Smith, K. Patel, H. Munford, A. Peet, D. A. Tennant, M. Jeeves, C. Ludwig, *Wellcome Open Res.* **2018**, *3*, 5. https://doi.org/10.12688/wellcomeopenres.13387.2

[16] S. Massou, C. Nicolas, F. Letisse, J.-C. Portais, *Phytochemistry* **2007**, *68*(16–18), 2330. https://doi.org/10.1016/j.phytochem.2007.03.011

[17] P. N. Reardon, C. L. Marean-Reardon, M. A. Bukovec, B. E. Coggins, N. G. Isern, *Anal. Chem.* **2016**, *88*(5), 2825. https://doi.org/10.1021/acs.analchem.5b04535

[18] M. P. Schätzlein, J. Becker, D. Schulze-Sünninghausen, A. Pineda-Lucena, J. R. Herance, B. Luy, *Anal. Bioanal. Chem.* **2018**, *410*(11), 2793. https://doi.org/10.1007/s00216-018-0961-6

[19] D. Schulze-Sünninghausen, J. Becker, M. R. M. Koos, B. Luy, *J. Magn. Reson.* **2017**, *281*, 151. https://doi.org/10.1016/j.jmr.2017.05.012

[20] P. Giraudeau, S. Massou, Y. Robin, E. Cahoreau, J.-C. Portais, S. Akoka, *Anal. Chem.* **2011**, *83*(8), 3112. https://doi.org/10.1021/ac200007p

[21] M. Mobli, J. C. Hoch, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 21. https://doi.org/10.1016/j.pnmrs.2014.09.002

[22] W. Willker, U. Flögel, D. Leibfritz, *J. Magn. Reson.* **1997**, *125* (1), 216. https://doi.org/10.1006/jmre.1996.1101

[23] M. Misiak, W. Koźmiński, K. Chmurski, K. Kazimierczuk, *Magn. Reson. Chem.* **2013**, *51*(2), 110. https://doi.org/10.1002/mrc.3917

[24] A. Le Guennec, J.-N. Dumez, P. Giraudeau, S. Caldarelli, *Magn. Reson. Chem.* **2015**, *53*(11), 913. https://doi.org/10.1002/mrc.4258

[25] G. Bodenhausen, D. J. Ruben, *Chem. Phys. Lett.* **1980**, *69* (1), 185. https://doi.org/10.1016/0009-2614(80)80041-8

[26] A. G. Palmer,; J. Cavanagh,; P. E. Wright,; M. Rance, *J. Magn. Reson. 1969* **1991**, *93* (1), 151. https://doi.org/10.1016/0022-2364(91)90036-S

[27] L. Kay, P. Keifer, T. Saarinen, *J. Am. Chem. Soc.* **1992**, *114*(26), 10663. https://doi.org/10.1021/ja00052a088

[28] J. Schleucher, M. Schwendinger, M. Sattler, P. Schmidt, O. Schedletzky, S. J. Glaser, O. W. SØrensen, C. Griesinger, *J. Biomol. NMR* **1994**, *4*(2). https://doi.org/10.1007/BF00175254

[29] M. Bostock, D. Nietlispach, *Concepts Magn. Reson. Part a* **2017**, *46A*(2), e21438. https://doi.org/10.1002/cmr.a.21438

[30] K. Kazimierczuk, V. Y. Orekhov, *Angew. Chem., Int. Ed.* **2011**, *50*(24), 5556. https://doi.org/10.1002/anie.201100370

[31] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*(3). https://doi.org/10.1007/BF00197809

[32] V. Y. Orekhov, V. A. Jaravine, *Prog. Nucl. Magn. Reson. Spectrosc.* **2011**, *59*(3), 271. https://doi.org/10.1016/j.pnmrs.2011.02.002

[33] S. G. Hyberts, A. G. Milbradt, A. B. Wagner, H. Arthanari, G. Wagner, *J. Biomol. NMR* **2012**, *52*(4), 315. https://doi.org/10.1007/s10858-012-9611-z

[34] S. G. Hyberts, H. Arthanari, S. A. Robson, G. Wagner, *J. Magn. Reson. San Diego Calif 1997* **2014**, *241*, 60. https://doi.org/10.1016/j.jmr.2013.11.014

[35] J. C. J. Barna, E. D. Laue, *J. Magn. Reson. 1969* **1987**, *75* (2), 384. https://doi.org/10.1016/0022-2364(87)90047-3.
[36] S. G. Hyberts, K. Takeuchi, G. Wagner, *J. Am. Chem. Soc.* **2010**, *132*(7), 2145. https://doi.org/10.1021/ja908004w
[37] F. Delaglio, G. S. Walker, K. A. Farley, R. Sharma, J. C. Hoch, L. W. Arbogast, R. G. Brinson, J. P. Marino, *Am. Pharm. Rev.* **2017**, *20*(4).
[38] K. Golman, R. Zandt, in't; M. Thaning, *Proc. Natl. Acad. Sci. U. S. a.* **2006**, *103* (30), 11270. https://doi.org/10.1073/pnas.0601319103
[39] J. Kurhanewicz, D. B. Vigneron, K. Brindle, E. Y. Chekmenev, A. Comment, C. H. Cunningham, R. J. DeBerardinis, G. G. Green, M. O. Leach, S. S. Rajan, R. R. Rizi, B. D. Ross, W. S. Warren, C. R. Malloy, *Neoplasia* **2011**, *13*(2), 81. https://doi.org/10.1593/neo.101102
[40] K. M. Koczula, C. Ludwig, R. Hayden, L. Cronin, G. Pratt, H. Parry, D. Tennant, M. Drayson, C. M. Bunce, F. L. Khanim, U. L. Gunther, *Leukemia* **2016**, *30*(1), 65. https://doi.org/10.1038/leu.2015.187
[41] R. Dass, K. Grudziaz, T. Ishikawa, M. Nowakowski, R. Dębowska, K. Kazimierczuk, *Front. Microbiol* **2017**, *8*(1306). https://doi.org/10.3389/fmicb.2017.01306
[42] M. A. C. Reed, J. Roberts, P. Gierth, Ē. Kupče, U. L. Günther, *ChemBioChem* **2019**, *20*(17), 2207. https://doi.org/10.1002/cbic.201900084
[43] I. Alshamleh, N. Krause, C. Richter, N. Kurrle, H. Serve, U. L. Günther, H. Schwalbe, *Angew. Chem., Int. Ed.* **2020**, *59* (6), 2304. https://doi.org/10.1002/anie.201912919
[44] C. A. Sellick, R. Hansen, G. M. Stephens, R. Goodacre, A. J. Dickson, *Nat. Protoc.* **2011**, *6*(8), 1241. https://doi.org/10.1038/nprot.2011.366
[45] R. M. Higashi, T. W.-M. Fan, P. K. Lorkiewicz, H. N. B. Moseley, A. N. Lane, In *Mass Spectrometry in Metabolomics*; D. Raftery, Ed.; Springer New York: New York, NY, **2014**; Vol. *1198*, pp 147.
[46] B. Christensen, J. Nielsen, *Metab. Eng.* **1999**, *1*(4), 282. https://doi.org/10.1006/mben.1999.0117

How to cite this article: Jeeves M, Roberts J, Ludwig C. Optimised collection of non-uniformly sampled 2D-HSQC NMR spectra for use in metabolic flux analysis. *Magn Reson Chem*. 2020; 1–13. https://doi.org/10.1002/mrc.5089