Journal of Visualized Experiments

Pure Shift Nuclear Magnetic Resonance: a new tool for plant metabolomics.

--Manuscript Draft--

Article Type:	Invited Methods Collection - Author Produced Video
Manuscript Number:	JoVE62719R1
Full Title:	Pure Shift Nuclear Magnetic Resonance: a new tool for plant metabolomics.
Corresponding Author:	Juan Lopez PUCP: Pontificia Universidad Catolica del Peru Lima, Lima PERU
Corresponding Author's Institution:	PUCP: Pontificia Universidad Catolica del Peru
Corresponding Author E-Mail:	juan.lopez@pucp.edu.pe
Order of Authors:	Juan Lopez
	Vanessa Leyva
	Helena Maruenda
Additional Information:	
Question	**Response**
Please specify the section of the submitted manuscript.	Biochemistry
Please indicate whether this article will be Standard Access or Open Access.	Standard Access ($1400)
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement
Please provide any comments to the journal here.	
Please confirm that you have read and agree to the terms and conditions of the video release that applies below:	I agree to the Video Release

TITLE:

Pure Shift Nuclear Magnetic Resonance: a New Tool for Plant Metabolomics

AUTHORS AND AFFILIATIONS

Juan M. Lopez[1], Vanessa Leyva[1], Helena Maruenda[1]

[1]Departamento de Ciencias – Química, CERMN, Pontificia Universidad Católica del Perú, Lima, Peru

vleyva@pucp.edu.pe

Correspondence to:
Juan M. Lopez at juan.lopez@pucp.edu.pe
Helena Maruenda at hmaruen@pucp.edu.pe

SUMMARY

This paper presents the use of PSYCHE and SAPPHIRE-PSYCHE in the metabolic profiling of plants and includes detailed procedures for sample preparation and optimal Pure Shift NMR spectra recording. Examples through which the gain in resolution achieved by homonuclear decoupling allows a more comprehensive understanding of the system are discussed.

ABSTRACT

Nuclear Magnetic Resonance (NMR) is one of the most powerful tools used in metabolomics. It stands as a highly accurate and reproducible method that not only provides quantitative data but also permits structural identification of the metabolites present in complex mixtures.

Metabolic profiling by ^{1}H NMR has proven useful in the study of various types of plant scenarios, which include the evaluation of crops conditions, harvest and post- harvest treatments, metabolic phenotyping, metabolic pathways, gene regulation, identification of biomarkers, chemotaxonomy, quality control, denomination of origin, among others. However, signal overlapping of the large number of resonances with expanded *J*-coupling multiplicities complicates the spectra analysis and its interpretation, and represents a limitation for classical ^{1}H NMR profiling.

In the last decade, novel NMR broadband homonuclear decoupling techniques through which multiplet signals collapse into single resonance lines - commonly called Pure Shift methods - have been developed to overcome the spectra resolution problem inherent to ^{1}H NMR classical spectra.

Here a step-by-step protocol of the plant extract preparation and the procedure to record optimal Pure Shift PSYCHE and SAPPHIRE-PSYCHE spectra in three different plant matrices - *Vanilla* plant leaves, potato tubers (*S. tuberosum*), and Cape gooseberries (*P. peruviana*) - is presented. The effect of the gain in resolution in metabolic identification, correlation analysis and multivariate analyses, as compared against classical spectra, is discussed.

INTRODUCTION

The complete set of metabolites that comprise an organism - substrates, intermediates, and end products of biological processes - was coined in 1998 with the term, metabolome. It is well known that the metabolome is closely related to the phenotype, and it is of particular interest in plants as it reflects the direct interaction between the genotype and the environment[1, 2]. Hence, the characterization of the metabolomic profile has become of paramount importance in plants. Through the identification and quantification of biomarkers (key metabolites) and metabolic patterns, the discrimination between species, cultivars, development stages, pathogenic diseases, or environmental conditions (daily and seasonal changes, soils, water stress, mechanical stress, harvest and post- harvest treatments), among others, has been possible[3–5].

Mass spectrometry (MS) and Nuclear Magnetic Resonance (NMR) spectroscopy are the most widely used analytical platforms for this purpose. Contrary to MS methodologies, NMR stands as a highly reproducible, non-biased, quantitative, accurate, and non-destructive technique that requires minimal sample preparation, making it suitable for metabolomics studies. However, when compared to MS methods, the inherent low sensitivity is a limitation. In recent years and through the use of high-field magnets, cryogenic probes, micro-coil devices, and Dynamic Nuclear Polarization (DNP) methods, the sensitivity of NMR has been greatly improved. In the case of the latter approach, for instance, the sensitivity gain was in the level of two to three orders of magnitude[6, 7]. To date, almost 20% of the published metabolomics studies are NMR-based and the number is rising[7].

Even though Proton NMR is the most popular and sensitive experiment for NMR metabolomics fingerprinting, it has some drawbacks. First, all the ^{1}H NMR signals detected in the sample are distributed in a small window corresponding to the proton chemical shift window, which results in crowded spectra. Second, the homonuclear scalar coupling splits the signals into multiple components (signal multiplicity), spreading the proton signal over a wider frequency range, complicating furthermore the spectra reading by increasing crowding and signal overlapping. In addition, NMR metabolomics is employed in the analysis of mixtures usually containing 50 to 300 molecules at an NMR observable concentration, generating complex spectra comprised of 200 to 2000 peaks.

Homonuclear decoupling proton NMR, also known as Pure Shift, is a method that induces the collapse of a multiplet signal into a single peak. It stands as an excellent tool for increasing signal resolution in crowded spectra[8–10] and therefore represents a convenient tool for plants metabolomics[11].

In the last decade, new Pure Shift pulse sequences, increasing both sensitivity and decoupling performance, have emerged. Their range of applications have also expanded, from molecular structure elucidation[12, 13], to fluxomics[14], mixture assignment[15–17], translational diffusion measurements[18], enantiomeric discrimination[19], unit distribution in co-polymers[20], among others.

Historically, Broadband Pure Shift experiments suffer from low sensitivity and complicated processing methods, limiting their scope in the assessment of biological extracts[8]. In 2014, Foroozandeh et al. published a new Pure Shift experiment, PSYCHE (Pure Shift Yielded by Chirp Excitation), based on anti-z-COSY pulse sequence which yielded excellent homonuclear decoupling and improved sensitivity values[21]. However, as PSYCHE is a 2D interferogram experiment where chunks of time domain data are acquired, it suffers from periodic sideband artifacts that result from *J*-coupling modulation distortions at the edges of the chunk. In complex mixtures, these artifacts yield signals larger than those associated with metabolites present at very low concentrations, hindering the analysis[11]. There are two methods to remove these artifacts - TSE-PHYCHE[22] and a more recent modification of the PSYCHE experiment called SAPPHIRE-PSYCHE (Sideband Averaging by Periodic PHase Incrementation of Residual J Evolution) [23].

In 2019, we demonstrated for the first time[11] that the SAPPHIRE-PSYCHE Pure Shift method, which removes artifacts with almost no sensitivity penalty[23], could be employed for the analysis of complex biological mixtures, such as extracts of the fruits of *Physalis peruviana*, commonly known as Cape gooseberries[11]. We showed that these methods increase the performance of metabolomics data analyses such as metabolic assignment, correlation analysis and multivariate coefficients analysis[11]. Since then, several Pure Shift metabolomics studies on different biological matrices, such as soft corals[24], hypericum plants[25], honey[26, 27], tea[27], peppermint oil[26], and walnuts[28] have been addressed, demonstrating its importance as a new tool for metabolomics analysis. Paradoxically, the vast majority of these studies employed the standard and easy to implement PSYCHE pulse sequence, available from any spectrometer library, instead of the SAPPHIRE-PSYCHE pulse sequence, which has been shown to perform better. However, it requires better understanding of the pulse sequence for proper setup.

This paper is intended to help new users to apply Pure Shift methods in the study of plants, in particular, leaves of *Vanilla sp* (*V. planifolia* and *V. pompona*) [29], potato tubers (*S. tuberosum*) [30], and Cape gooseberries (*P. peruviana*) [31]. Sample preparation, NMR experimental set up, data acquisition, and data analysis are described in detail. Moreover, the protocol includes key notes to help researchers, new to the field, to properly set up PSYCHE and SAPPHIRE-PSYCHE experiments in the metabolomic profiling of plants.

PROTOCOLS

1. Sample preparation

1.1. Cape gooseberries

1.1.1. Place 100-200 g of freshly peeled fruits in a blender vase. Keep at 4 °C for 30 min, and then homogenize in a cold laboratory blender.

1.1.2. Immediately, transfer the juice to 50 mL plastic tubes, freeze them in liquid nitrogen and lyophilize to dryness for 4 to 5 days.

1.1.3. Grind the lyophilized material to a fine powder using an electric grinder.

NOTE: Handling of the dry material needs to be done quickly because the powder is highly hygroscopic.

1.1.4. Weigh 1 g of the ground material and add 10 mL of ultrapure water. Vortex for 1 min.

1.1.5. Sonicate for 20 min at 10 °C, and then centrifuge at 23,000 × *g* for 20 min at 10 °C.

1.1.6. Recover the supernatant and filter it through a 13 mm polytetrafluoroethylene (PTFE) 0.45 µm syringe filter.

1.1.7. Lyophilize 1 mL of the filtrated extract to dryness and then re-suspend the obtained solid in 0.9 mL of 200 mM sodium oxalate buffer pH 4. Vortex.

1.1.8. Lyophilize the resulting sample to dryness and dissolve it in 0.9 mL of deuterium oxide containing 5 mM of 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (TMSP-d_4).

1.1.9. Fill the NMR tube with 0.6 mL of the sample using a micropipette.

1.2. *Vanilla* leaves

1.2.1. Collect the leaves, clean them with damp tissue paper and freeze them whole in liquid nitrogen.

1.2.2. Break the leaves into small pieces and lyophilize for 4 days until dryness.

1.2.3. Grind the dry matter to a fine powder using an electric grinder.

1.2.4. Weigh 50 mg of ground material and add 0.75 mL phosphate buffer pH 6.0 in deuterium oxide containing 0.1% of TMSP (w/w) and 0.75 mL methanol-d4. Vortex for 1 min.

1.2.5. Sonicate for 20 min at 25 °C.

1.2.6. Centrifuge at 13,000 × *g* for 10 min at 25 °C.

1.2.7. Recover the supernatant (~1.3-1.4 mL) and filter it through a 13 mm PTFE 0.45 µm syringe filter.

1.2.8. Fill the NMR tube with 0.6 mL of the filtered sample using a micropipette.

1.3. Potato tubers

1.3.1. Peel and slice 4 to 8 tubers. Immediately, place approximately 125 g of material in stand-up bags and freeze them in liquid nitrogen.

NOTE: To avoid oxidation during handling keep the potato damp.

1.3.2. Lyophilize for 4 to 6 days until complete dryness.

1.3.3. Grind the dry matter to a fine powder using an electric grinder.

1.3.4. Weigh 160 mg of ground tuber and add 1.6 mL of deionized water. Vortex for 1 min.

1.3.5. Sonicate for 45 min at 10 °C.

1.3.6. Centrifuge at 23,000 x *g* for 20 min at 10 °C.

1.3.7. Recover the supernatant (~1.5 -1.6 mL) and evaporate it to dryness in a refrigerated centrifugal vacuum concentrator for 16 h, at 10 °C.

1.3.8. Add to the solid obtained (20-25 mg) 0.9 mL of 100 mM sodium oxalate buffer pH 4, vortex, and evaporate for 16 hours at 10 °C.

1.3.9. Dissolve the solid obtained in 0.9 mL of deuterium oxide containing 3 mM of TMSP.

1.3.10. Centrifuge at 23,000 x g for 5 min at 10 °C and filter the supernatant directly into the NMR tube through a 13 mm PTFE 0.45 µm syringe filter.

NOTE: In this case, direct filtration into the NMR tube was performed to diminish steps in the preparation of more than 1000 samples.

2. NMR Data Acquisition and Processing

2.1. NMR initial setup

2.1.1. Transfer the samples to the NMR spectrometer.

2.1.2. Tune and match the probehead.

2.1.3. Lock and shim the sample.

2.1.4. Calibrate the 90° hard pulse. Calibrate the 90° pulse using any of the standard procedures.

2.1.5. Run a standard 1D proton NMR spectrum.

2.2. PSYCHE experiment

2.2.1. Select the reset_psyche_1d pulse sequence from the Bruker Topspin library (**Figure S1**). Use the following standard parameters: 5 kHz spectral width (SW2), at least 1 or 2 seconds of relaxation recovery delay (D1), 16 dummy scans (DS), 64 or 128 complex data points per block (L31), and 64 or 128 scans (NS) (**Figure S1**).

NOTE: L31 is the number of complex digital points acquired in each Pure Shift block, best to be set to a power of 2.[21]

2.2.2. Set the desired CHIRP pulse flip angle excitation (CNST61) and 10 kHz for the CHIRP pulse bandwidth (CNST60) (**Figure S2**).

NOTE: The PSYCHE experiment is based on an anti-z-COSY scheme; consequently the CHIRP pulse flip angle needs to be small to avoid recoupling artifacts (**Figure 1**). The absolute intensity increases with the excitation flip angle. The periodic artifacts are also enhanced, spreading into the spectrum and increasing the "noise" (**Figure 1**). The "noise" becomes a combination of standard noise and chuncking artifacts. A good compromise between sensitivity and low recoupling artifacts is to set CNST61 = 20°.[19, 22]

2.2.3. Set the hard pulse length (P1) to the previously calibrated value and the PSYCHE shape pulse length to 30 ms (P49) (**Figure S2**).

NOTE: It is very important to calibrate the hard pulse value as the shape pulse powers will be automatically calculated from this value.

2.2.4. Choose the Crp_psyche.20 (SPNAM 37) shape pulse for the PSYCHE element (**Figure S2**).

2.2.5. Set the strength of the pulse field gradient applied during the PSYCHE element (GPZ0). Choose RECT.1 for the gradient shape pulse (GPNAM 0) (**Figure S2**).

NOTE: A weak magnetic field gradient is applied during the PSYCHE element, normally, between 1% to 4% of the maximum strength of the gradient, depending on the probe.

2.2.6. Set the number of blocks to acquire in order to reconstruct the Pure Shift FID (TD1) (**Figure S3**).

NOTE: PSYCHE is acquired as a pseudo-2D experiment where TD1 is the number of Pure Shift interferogram blocks. The spectrum resolution depends on the size of the spectral window (SW1) and the total number of acquired points, which is TD1*2*L31. Typically, 16 or 32 blocks with 64 or 128 complex points per block will provide enough digital resolution. As PSYCHE is recorded in an interferogram manner, a higher number of blocks increase the digital resolution, but, also the total acquisition time[19]. Homonuclear *J*-couplings evolve during each block resulting in an oscillating *J* modulation pattern[21, 23]. After Fourier transform, this generates

periodic sideband artifacts that depend on the length of the block (**Figure 2**). To reduce artifacts, the duration of the block must be short, typically less than 16 ms (block duration = 2*in0: **Figure S1**). If the block duration is high, reduce L31.

2.2.7. Process the data with Bruker's Proc_reset AU program and Fourier transform.

NOTE: We recommend to transform the spectrum using zero filling and a sine bell apodization (**Figure S4**).

2.3. SAPPHIRE-PSYCHE experiment

2.3.1. Select the SAPPHIRE-PSYCHE pulse sequence and set the pulse sequence parameters. Standard parameters would be the following: 5 kHz spectral width (SW3), at least 1 or 2 seconds of relaxation delay (D1), 16 dummy scans (DS), 8 or 16 scans per increment (NS) and D2 to 14 ms (**Figure S5**).

NOTE: This sequence is not in Bruker's repertoire, however, the sequence and the processing programs may be obtained from the Manchester NMR Methodology Group website, (https://www.nmr.chemistry.manchester.ac.uk/?q=node/426) [23]. The D2 delay ensures that T2 relaxation remains constant with each J modulation increment. D2 needs to be greater than 1/4*SW1+p16+2*d16.[23]

2.3.2. Set the desired CHIRP pulse flip angle excitation (CNST20) and 10 kHz for the CHIRP pulse bandwidth (CNST21) (**Figure S6**).

NOTE: As in the regular PSYCHE experiment, the CHIRP pulse flip angle needs to be short to avoid recoupling artifacts. CNST20 = 20° is a good compromise between sensitivity and low recoupling artifacts[21, 23, 25].

2.3.3. Set the hard pulse length (P1) to the previously calibrated value and the PSYCHE shape pulse length to 30 ms (P40) (**Figure S6**).

NOTE: It is important to calibrate the hard pulse value as the shape pulse powers will be automatically calculated from it.

2.3.4. Choose the PSYCHE_Saltire_10kHz_30m shape pulse for the PSYCHE element (**Figure S6**).

2.3.5. Set the strength of the pulse field gradient applied during the PSYCHE element (GPZ10). Choose RECT.1 for the gradient shape pulse (GPNAM 10) (**Figure S7**).

NOTE: A weak magnetic field gradient is applied during the PSYCHE element, normally, between 1% and 4% of the maximum strength of the gradient, value that depends on the probe.

2.3.6. Set the number of SAPPHIRE J modulation increments in F2 (TD2) (**Figure S7**).

NOTE: normally 8 increments ensure an excellent suppression of sideband artifacts (**Figure 2 and 3**). The total number of scans of the final Pure Shift FID is NS*TD2.[23]

2.3.7. Set the F1 and F2 spectral windows (SW1 and SW2) (**Figure S5**).

NOTE: SW2=SW3/(2*TD2) and SW3/SW1 = TD2*N, were TD2 and N are even integers.[23] The SAPPHIRE-PSYCHE experiment is acquired as a pseudo 3D where F2 encodes the *J*-coupling artifact phase modulation and F1 the Pure Shift interferogram acquisition[20]. Since SAPPHIRE-PSYCHE removes *J* modulation sidebands, the interferogram Pure Shift block duration could be longer than regular PSYCHE (Pure Shift block duration = 1/SW1), typically between 20 to 40 ms (**Figure 2**). However, longer chunk data acquisition leads to higher *J*-coupling evolutions, which would require more *J*-coupling phase modulation increments to remove the stronger sidebands attained.[23]

2.3.8. Set the number of Pure Shift blocks (TD1) (**Figure S7**).

NOTE: Since SAPPHIRE-PSYCHE needs to compensate the *J*-coupling phase modulation of the first block, an extra block needs to be acquired. Typically, 17 (16+1) or 33 (32+1) blocks give enough digital resolution.[23]

2.3.9. Process the data executing the pm_pshift and the pm_fidadd AU programs followed by Fourier transform[23].

NOTE: We recommend to transform the spectrum using zero filling and a sine bell apodization (**Figure S4**).

REPRESENTATIVE RESULTS

NMR spectrum analysis

PSYCHE experiments increase spectra resolution by collapsing coupled resonances into singlets[21], which in turn reduces overlap and facilitates assignment and data analysis. Pure Shift NMR can be applied to plant extracts. Here we demonstrate its use in three different matrices: vanilla leaves, potato tubers, and *Physalis peruviana* fruits. The resolution enhancement achieved in the spectra of these plant extracts is clear from **Figures S8-S11**.

In order to assess the effect of these pulse sequences in the resolution of the NMR signal, the width of the line in frequency units at 10% of the maximum height, W_{10}, were calculated for several types of resonances, all exemplified in spectra obtained with *Vanilla sp* extracts - Pure Shift against classical ^{1}H NMR - **Figure 4**. Overall, the expansion of coupled resonances reached W_{10} values from 1 to 60 Hz, whereas singlet peaks varied between 1 to 10 Hz: the methyl group of acetic acid (1.97 ppm in **Figure 4**) attained a W_{10} value of 2.0 Hz in classical ^{1}H-NMR and 2.1 Hz in the SAPPHIRE-PSYCHE (S) spectrum. The anomeric resonance of sucrose (at 5.40 ppm), a doublet with a *J*-coupling of 3.9 Hz, extends over a region equivalent to a W_{10} of 6.5 Hz, with each

individual peak accounting for 2.6 Hz of W_{10} (**Figure 4**). This value was higher than that of the collapsed peak obtained in the SAPPHIRE-PSYCHE spectrum, W_{10} = 1.9 Hz (**Figure 4**). In the case of the malic acid Hβ´ (at 2.55 ppm), a doublet of doublets with *J* =7.8 Hz, 15.6 Hz, extends over 29.4 Hz (**Figure 4**), with calculated W_{10} values for each individual peak between 4.7 and 4.9 Hz. This multiplet signal collapsed by Pure Shift into a single line (at 2.55 ppm), expanding over a W_{10} value of 4.7 Hz (**Figure 4**).

Signal multiplicity becomes more complex in highly coupled hydrogens where the constitutive peaks of the multiplet are no longer easily distinguishable, forming an almost continuous signal. This is the case for the homocitric acid hydrogens, Hγ (at 2.24 ppm), Hδ (at 2.02 ppm), and Hδ´ (at 1.91 ppm), Figure 4. The Hγ, Hδ, and Hδ´multiplets with W_{10}=38.3 Hz, W_{10}=41.5 Hz, and W_{10}=34.8 Hz, respectively, collapsed to singlets of W_{10}=7.8 Hz (Hγ, 2.24 ppm), W_{10}=8.0 Hz (Hδ, 2.02 ppm) and W_{10}=6.2 Hz (Hδ´, 1.91 ppm), Figure 4. This enhanced resolution allowed better discrimination of other overlapped signals corresponding to malic Acid (Hβ, 2.77 ppm Hβ´, 2.55 ppm), homocitric acid (Hα 2.80 ppm and Hα´, 2.63 ppm), and homocitric acid lactone (Hα´, 2.84 ppm and Hδ 2.65 ppm) (**Figure 4**).

The improvement in resolution in the SAPPHIRE-PSYCHE homonuclear decoupled spectra that permits better discrimination among signals is further demonstrated in two highly compromised regions, 4.85 – 5.08 ppm and 7.05 – 7.31 ppm, where important metabolites involved in vanilla fragrance metabolic pathway could be identified: glucoside A (see structure in Figure 5), CH-2´,6´ (at 7.28 ppm), CH-2",6" (at 7.18 ppm), CH-3´,5´ (at 7.09 ppm), CH-3",5" (at 7.07 ppm), CH_2-7´ (at 4.97 ppm and 5.05 ppm), CH_2-7'' (at 4.89 ppm), CH-Glc (at 4.96 ppm) and CH'-Glc (at 4.93 ppm), glucoside B (see structure in Figure 4), CH-2´,6´ (at 7.29 ppm), CH-2",6" (at 7.23 ppm), CH-3´,5´ (at 7.09 ppm), CH-3",5" (at 7.07 ppm), CH_2-7´ (at 4.99 ppm), CH_2-7'' (at 4.91 ppm), hydroxybenzyl alcohol (7.23 ppm) and hydroxybenzyl alcohol glucoside (7.10 and 4.51 ppm), **Figure 4**.

The same results were observed with Cape gooseberry extracts and in the potato tuber analysis. In these two cases, several crowded regions on regular ^{1}H-NMR, PSYCHE, and SAPPHIRE-PSYCHE have been expanded for comparisons (**Figures 6 and 7**).

It is clear that PSYCHE and SAPPHIRE-PSYCHE experiments clearly improved signal resolution of Cape gooseberry extracts for sucrose (4.04 ppm), β-fructose (4.02 ppm and 3.99 ppm), β-glucose (3.24 ppm), proline (2.34 ppm, 2.07 ppm, and 2.00 ppm), glutamic acid (2.16 ppm), and glutamine (2.13 ppm) (Figure 6), as we showed in previous work,[11] as well as it did in the case of the *S. tuberosum* extracts: GABA (1.91 ppm, 2.33 ppm, and 3.02 ppm), pyroglutamic acid (2.04 ppm, 2.41 ppm, and 2.51 ppm), proline (2.00 ppm, 2.07 ppm, 3.33 ppm, and 3.42 ppm), glutamine (2.14 ppm and 2.46 ppm), valine (2.27 ppm), citric acid (2.64 ppm and 2.76 ppm), and β-glucose (3.25 ppm and 3.47 ppm), **Figure 7**.

The PSYCHE sequence is relatively easy to use and it has been successfully implemented in a wide range of applications[9]. To attain the decoupled spectrum, the pulse sequence acquires small chunks of FID refocussing the *J*-coupling at the middle of each block. However, a small *J*-coupling evolution occurs during each block and generates the periodic sidebands artifacts which normally

represent less than 5 % of their parent peak[21]. The presence of these *J* modulation artifacts, inherent to PSYCHE experiments,[21] are evident in **Figure 6** expanded zones. In the analysis of pure compounds these artifacts can be neglected. In biological samples, this may not be the case, because, as shown in **Figure 6** for proline (4.13 ppm, 3.41 ppm, and 3.33 ppm), asparagine (3.97 ppm), myo-inositol (3.27 ppm), GABA (3.03 ppm), and malic acid (2.67 ppm), and in **Figure 6** for pyroglutamic acid (2.04 ppm, 2.41 ppm, and 2.51 ppm) and valine (2.27 ppm), metabolites at higher concentrations generate artifacts as large as some of the signals belonging to compounds present at low concentrations compromising in this way the accuracy of the metabolic profiling.

The SAPPHIRE-PSYCHE experiment is a modification of the regular PSYCHE sequence in which these periodic artifacts are removed by systematic phase modulation, achieved by shifting the *J* refocusing point.[23] Consequently, the SAPPHIRE-PSYCHE experiment allows to ensure a much cleaner Pure Shift spectrum, as shown in Figure 6 for proline (4.13 ppm, 3.41 ppm, and 3.33 ppm), asparagine (3.97 ppm), myo-inositol (3.27 ppm), GABA (3.03 ppm), and malic acid (2.67 ppm), in the case of cape gooseberries,[11] and in Figure 7 for pyroglutamic acid (2.04 ppm, 2.41 ppm, and 2.51 ppm) and valine (2.27 ppm), in the case of potatoes.

Another artifact from which all Pure Shift experiments suffer are those generated by the strong coupling effect. Some important primary metabolites such as citric acid (2.64 ppm and 2.76 ppm) and glutamine (2.14 ppm and 2.46 ppm), exhibit strong coupling artifacts, as shown in potato tuber extract (**Figure 7**). To date, there is no pulse sequence that can conveniently eliminate this problem; however, SAPPHIRE-PSYCHE performs better than regular PSYCHE[8,9,23].

Correlation matrix analysis

One of the main advantages of NMR spectroscopy is that the relative concentrations of metabolites in a mixture are directly proportional to the intensities of their signals. Then, a metabolite pairwise correlation matrix may be obtained directly from the spectra correlation matrix[32].

The ^{1}H-NMR correlation matrix, commonly known as STOCSY[32] (Statistical Total Correlation Spectroscopy), is usually represented as a pseudo-2D spectrum, where each cross peak is a correlation coefficient between two signals (**Figure 8**) [11]. STOCSY displays high correlation between signals that belong to the same molecule, but also with signals from molecules that pertain to the same metabolic pathway[11, 32]. Hence, correlations patterns provide information about the physiological state of the system, and can therefore be employed as a physiological stage fingerprint.[33]

The main limitation of STOCSY is signal overlapping, which decreases the pairwise correlation[11, 32]. Moreover, *J*-coupling multiplicity correlations, lead to highly complex patterns, complicating further the analysis[11, 32]. The use of SAPPHIRE-PSYCHE STOCSY, shown in Figure 8, enhances correlation values because it reduces *J*-coupling multiplicity correlations into a single peak, thereby reducing the overlapping[11].

Several regions have been expanded in **Figure 8**. Malic acid shows complex *J* multiplet patterns

associated with Hα (4.41 ppm), Hβ (2.82 ppm), and Hβ' (2.67 ppm); in Pure Shift STOCSY those signals collapse into single correlation peaks (**Figure 9A**).[11] Same results are observed for β-glucose were intermolecular and intramolecular (α-glucose, α-fructose, and β-fructose) correlations are better depicted by SAPPHIRE-PSYCHE (**Figure 9B**).[11] Some amino acids also display strong intramolecular correlations in Cape gooseberry extracts. However, their identification in regular ^{1}H-NMR STOCSY is compromised by the overlap in crowded regions. With Pure Shift STOCSY, proline, alanine, glutamine, and glutamic acid intermolecular and intramolecular correlations are better depicted (**Figures 9C and 9D**) [11].

Multivariate analysis

Multivariate analysis is one of the main tools employed in addressing metabolomics data[34, 35]. While sample discrimination by PCA (Principal component analysis) or PLS-DA (Partial least squares-discriminant analysis) could easily be achieved through regular ^{1}H-NMR spectra, the interpretation of the loading is better addressed through pure-shift data[11, 24, 25, 27, 28].

In **Figure 10**, we show the PLS-DA score plot obtained using SAPPHIRE-PSYCHE (Figure 10A) and regular proton spectra (**Figure 10B**) in the discrimination of six ecotypes of *P. peruviana* extracts[11]. Even though there are some studies which claim better discrimination performance when using Pure Shift data[24, 25], our results show that the performance was barely affected by the homonuclear decoupling[11]. In the case of the analysis of the loadings data, as shown in **Figures 10C and 10D**, the increased resolution attained with the SAPPHIRE-PSYCHE data simplified the analysis and allowed a better identification of the specific metabolites responsible of *P. peruviana* discrimination, namely, α-glucose, β-glucose, α-fructose, β-fructose, sucrose, citric acid, and alanine[11]. This resolution gain was also critical when combining PLS coefficients analysis with STOCSY correlation, **Figures 10C and 10D**[11]. The strong correlation between α-glucose (STOCSY vector at 5.23 ppm is color-coded on the first PLS component) and β- glucose, α-fructose and β- fructose – all metabolites which evolve from the same metabolic pathway – and its anti-correlation with respect to sucrose, are evident[11]. With normal STOCSY analysis, the extensive overlap did not allow clear depiction among correlation coefficients and produce the loss of this particular metabolic pathway information (**Figure 10**) [11].

FIGURES:

Figure 1. (**A**) PSYCHE spectra of *Vanilla planifolia* extracts using different CHIRP pulse flip angles for the PSYCHE element: left, flip angle values; right, intensity multiplication factor. (**B**) and (**C**) Graphs show the intensities and signal-to-noise of five decoupled peaks (black - 4.31 ppm, orange - 4.16 ppm, red - 2.65 ppm, yellow - 2.63 ppm, and blue 2.56 ppm) as a function of the flip angle of the PSYCHE CHIRP pulse, respectively. The signal-to-noise ratio (S/N) (more precisely, signal to noise + artifact ratio) was calculated using the signal of maximum intensity against that of the noise, value calculated over a 2ppm range: from 7.75 ppm to 9.75 ppm (**Figure S12**).

Figure 2. PSYCHE (**A**) and SAPPHIRE-PSYCHE (**B**) spectra of *Vanilla planifolia* extracts with different interferogram Pure Shift block duration. The experiments were acquired in order to maintain the same digital resolution and sensitivity. PSYCHE (**A**) parameter from bottom to top: 128 scans, 64 interferogram blocks, 6.4 ms block length, total acquisition time 6h07min; 128

scans, 32 interferogram blocks, 12.8 ms block length, total acquisition time 3h04min; 128 scans, 16 interferogram blocks, 25.6 ms block length, total acquisition time 1h32min; 128 scans, 8 interferogram blocks, 51.2 ms block length, total acquisition time 46min; 128 scans, 4 interferogram blocks, 102.4 ms block length, total acquisition time 24min. SAPPHIRE-PSYCHE (**B**) parameter from bottom to top: 16 scans, 8 *J* modulation increments, 65 interferogram blocks, 6.4 ms block length, total acquisition time 6h19min; 16 scans, 8 *J* modulation increments, 33 interferogram blocks, 12.8 ms block length, total acquisition time 3h13min; 16 scans, 8 *J* modulation increments, 17 interferogram blocks, 25.6 ms block length, total acquisition time 1h40min; 16 scans, 8 *J* modulation increments, 9 interferogram blocks, 51.2 ms block length, total acquisition time 53min; 16 scans, 8 *J* modulation increments, 5 interferogram blocks, 102.4 ms block length, total acquisition time 30min.

Figure 3. (**A**) spectra of *Vanilla planifolia* extracts with 102.4 ms chunk length: top, PSYCHE 128 scans, 4 Pure Shift increments; middle top, 4 scans, 4 Pure Shift increments, 32 SAPPHIRE increments; middle bottom, 8 scans, 4 Pure Shift increments, 16 SAPPHIRE increments; bottom, 16 scans, 4 Pure Shift increments, 16 SAPPHIRE increments. (**B**) spectra of *Vanilla planifolia* extracts with 51.2 ms chunk length: top, PSYCHE 128 scans, 8 Pure Shift increments; middle top, 4 scans, 9 Pure Shift increments, 32 SAPPHIRE increments; middle bottom, 8 scans, 9 Pure Shift increments, 16 SAPPHIRE increments; bottom, 16 scans, 9 Pure Shift increments, 16 SAPPHIRE increments.

Figure 4. Selected expansion regions of ^{1}H NMR (1H) and SAPPHIRE-PSYCHE (S) spectra *V. planifolia* (1.85 – 2.9 ppm) and *V. pompona* (4.85 – 7.31 ppm), showing signal assignments.

Figure 5. Structures of vanilla fragrance precursors: 4-hydroxybenzyl alcohol (4-HBA), 4-HBA glucoside, glucoside A, and glucoside B.

Figure 6. Selected expansion regions of ^{1}H NMR (1H), PSYCHE (P), and SAPPHIRE (S) spectra of an aqueous extract of Cape gooseberry (Bambamarca I Peruvian Andean region[28]) showing signal assignments (Reprinted with permission from Lopez et al.[11]).

Figure 7. Selected expansion regions of ^{1}H NMR (1H), PSYCHE (P), and SAPPHIRE (S) spectra of an aqueous extract of potato showing signal assignment**.**

Figure 8. Selected expanded regions (3.20 ppm–4.30 ppm) of two-dimensional STOCSY NMR spectra obtained with data from six Cape gooseberry extracts showing correlation values (r^2) above 0.85: (**A**) regular ^{1}H NMR STOCSY and (**B**) SAPPHIRE-PSYCHE STOCSY. Reprinted with permission from Lopez et al.[11]

Figure 9. STOCSY representations of NMR spectra of six different Cape gooseberry extracts showing correlations (r^2), in the left without homodecoupling and in the right with homodecoupling for regions: (**A**) 4.38–4.42 ppm and 2.40–4.42 ppm with r^2 above 0.80 for malic acid (MA) signal (Hα-MA); (**B**) 3.21–3.27 ppm and 3.21–4.67 ppm with r^2 above 0.85 for β-glucose signal (H2-β-Gluc); (**C**) 2.30–2.38 ppm and 1.25–4.36 with r^2 above 0.93 for proline (Pro) signal

(Hβ'-Pro); (**D**) 2.15–2.17 ppm and 1.25–4.5 ppm with r^2 above 0.90 for glutamic acid (Glu) signal (Hβ-Glu); α-glucose, α-fructose, β-fructose are symbolized as α-Gluc, α-Fruc, and β-Fruc, respectively. Reprinted with permission from Lopez et al.[11]

Figure 10. PLS scores plot of Cape gooseberries extracts grown in six different Andean regions[11, 28] (San Marcos: red circles, Celendin III: brown triangles, Bambamarca I: blue stars, Celendin I: yellow triangles, Bambamarca II: green squares, Celendin II: magenta diamonds) based on (A) classical ^{1}H NMR and (B) SAPPHIRE-PSYCHE experiments. Hotelling's T^2 ellipses were set to 95% confidence level. Combination of PLS1 loadings and 1D STOCSY for α-glucose correlation using the STOCSY signal at 5.23 ppm as driver peak. The coefficient of determinations (r^2) have been color coded and projected on the coefficients of the first PLS component: (A) 1D STOCSY obtained with SAPPHIRE-PSYCHE data (top) and its expansion 3.15–4.17 ppm (bottom); (B) 1D STOCSY obtained with ^{1}H NMR data (top) and its expansion 3.15–4.17 ppm (bottom); α-glucose, β-glucose, α-fructose, β-fructose, and sucrose are symbolized as α-G, β-G, α-F, β-F, and S, respectively. PLS-DA and STOCSY analysis was perform using MATLAB Version R2018a. (Reprinted with permission from Lopez et al.[11])

DISCUSSION

Metabolite structural identification and quantitation are key issues in the characterization of the metabolome, data that when subjected to multivariable analyses permits to better understand the biological system under study. Sample preparation and data acquisition are critical aspects that need optimization in order to provide reliable results.

In this article, we describe and illustrate the sample preparation for NMR analysis of three different plant matrices. As with any extraction procedure, the amount of solvent per gram of material and the physical properties of the selected solvent will determine the chemical composition of the final extract and the concentration of the extracted metabolites. In the case of NMR metabolomic profiling, pH, reproducibility between independent sample extractions, and final amount of extract in the NMR tube are also aspects that need optimization. The importance of reproducibility in metabolomics is to avoid the introduction of uncorrelated variance, which could lead to unreliable results. In our experience, optimal extraction conditions were attained with the dry and grinded plant material. In the case of Cape gooseberries, the dry product was very difficult to handle (highly hygroscopic) so the fresh berries were homogenized first, prior to lyophillization.

In the case of spectra acquisition, Pure Shift experimental setup is of particular importance, as wrong parameters can lead to chunking and recoupling artifacts. The theory behind the principles of Pure Shift experiments, extensively reviewed elsewhere[8–10], is important to understand how to correctly configure the pulse sequence and implement it as a routine experiment.

In brief, most Pure Shift experiments are based on refocusing the *J*-coupling evolution during chemical shift recording. This is typically accomplished by a *J*-coupling refocusing element that selectively inverts "passive" spins, while the "active" spins remain unaffected. PSYCHE and SAPPHIRE-PSYCHE are based on an anti-z-COSY experiment where "passive" spins are statistically

inversed.

The PSYCHE element, which consists of two swept-frequency low flip angle pulses in the presence of a weak pulsed field gradient, induces frequency spatiotemporal averaging, selecting the anti-diagonal COSY terms while suppressing zero quantum and cross correlation terms. Accordingly, to avoid the inherent recoupling artifacts, the CHIRP pulse flip angle needs to be short (**Figure 1**). Typically, a 20° flip angle is a good compromise between sensitivity and decoupling performance (**Figure 1**). Therefore, pulse calibration is critical for the quality and sensitivity of the spectrum.

Metabolomic studies are usually associated with spectra recording on a large number of samples, which implies that pulse calibration must be fast or automatic. In our experience, if the samples are prepared in exactly the same way, the hard pulse length variability among samples is less than ± 0.2 µs. We normally calibrate between 6 and 12 samples and then use the average value attained to set-up the whole group of samples. In the case that the pulse length variability from sample to sample were higher, automatic calibration of each sample should be performed using Topspin pulsecal automation program on Bruker spectrometer.

The second important parameter to consider is the length of the recorded blocks during the stepwise interferogram acquisition[21, 23]. Interferogram acquisition consists in recording the FID by small chunks, with the refocusing point of the *J*-coupling evolution always coinciding with the center of the acquired chunk. The decoupled FID is constructed by concatenating each successive chunk[8–10]. To ensure that the block acquisition does not truncate chemical shift evolution, the start of each data recording must exactly match the end of the previous chunk.

Although this procedure allows us to obtain a homodecoupled spectrum, the small *J*-coupling evolution during each block generates periodic sidebands artifacts directly dependent on the chunk length. On the other hand, the spectral digital resolution depends on the spectral window and on the total duration of the decoupled FID, which, in turn, depends on the block length and number of recorded blocks. Therefore, to reduce periodic artifacts without sacrificing resolution, the block duration should be short, and the total number of recorded blocks should be high. These conditions, however, will highly increase the total experimental acquisition time without increasing sensitivity (**Figure 2**). Typically, a PSYCHE experiment acquired with 16 to 32 blocks of 10 to 16 ms duration gives enough digital resolution in a reasonable experimental time (30 min to 5 h) (**Figure 2**).

In the case of SAPPHIRE-PSYCHE, an experiment that is acquired as a pseudo 3D, one of the indirect dimensions encodes for the Pure Shift interferogram acquisition and the other for the phase modulation of the periodic artifacts through the systematic shift of the *J* refocusing point in each chunk.

As periodic artifacts are strongly suppressed by SAPPHIRE, the block lengths could be longer; however, very long chunks strongly affect the intensity of the signals (**Figures 2 and 3**). In SAPPHIRE, *J* modulation increments contribute to spectrum sensitivity, therefore, the resulting decoupled FID total number of scans is equal to TD2 * NS (**Figures 2 and 3**)[23]. Generally speaking,

eight *J* modulation increments ensure an excellent periodic artifact suppression and more than eight increments have very little effect on the quality of the spectrum, even if long chunk lengths are used (**Figure 3**) [23]. Pure Shift increments of 33 or 17 with durations between 20 to 40 ms ensure a good spectral digital resolution.

One limitation of both of these Pure Shift pulse sequences, PSYCHE and SAPPHIRE-PSYCHE, is quantification by absolute metabolite-internal standard integration. In regular ^{1}H-NMR, the integrated intensity is directly proportional to the concentration of each metabolite. In PSYCHE, this is no longer the case, because a number of phenomena distort the signals and affect integration. For instance, the total integral value diminishes due to T2 relaxation during the pulse sequence spin selection. Also, the truncated *J* coupling evolution during the chunk acquisition which generates sidebands artifacts, disrupts the Lorentzian shape of the signal. Hence, the integral is now comprised by areas under the main peak and under all the sidebands, complicating signal integration[8, 9, 21, 23]. The frequency and magnitude of the sidebands are directly related to the chunk length but also to intrinsically molecular properties such as relaxation and the *J* coupling magnitude and multiplicity: higher *J* coupling magnitudes and higher multiplicities, lead to more distorted signals. In the case of SAPPHIRE, even though this NMR experiment efficiently removes the sideband artifacts, the signal intensities are compromised by the truncated *J* coupling evolution. The sum of each *J* modulated increment, generates an averaged decoupled FID where signal diminishment is directly related to the chunk length and the *J* coupling magnitude and multiplicity[23]. Moreover, CHIRP pulse flip angles generate recoupling artifact which also affects each signal differently, further complicating the quantification[21]. The magnitude of the effect of these pulse sequences in quantitative analysis was assessed in our earlier Cape gooseberry study yielding errors of around 10% to 30%[11].

Finally, we can conclude that Pure Shift is an excellent new tool for plant metabolomics, it drastically increases the spectrum resolution, allowing a finer correlation matrix analysis and a better interpretation of multivariate analyzes[11, 24, 25, 27, 28].

ACKNOWLEDGMENTS

This study was funded by the Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC) - Programa Atracción de Investigadores Cienciactiva – Contract # 008-2017-FONDECYT.

DISCLOSURES

The authors have no conflicts of interest to declare.

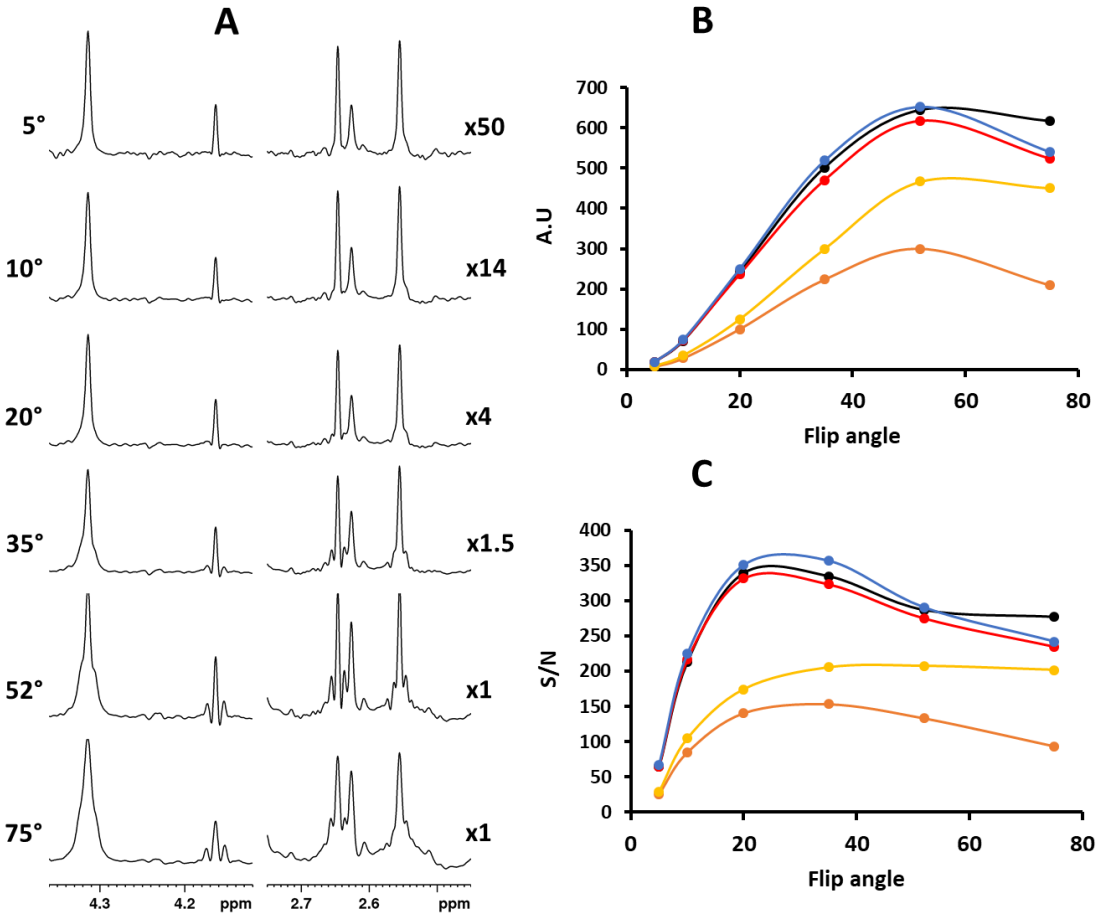
REFERENCES

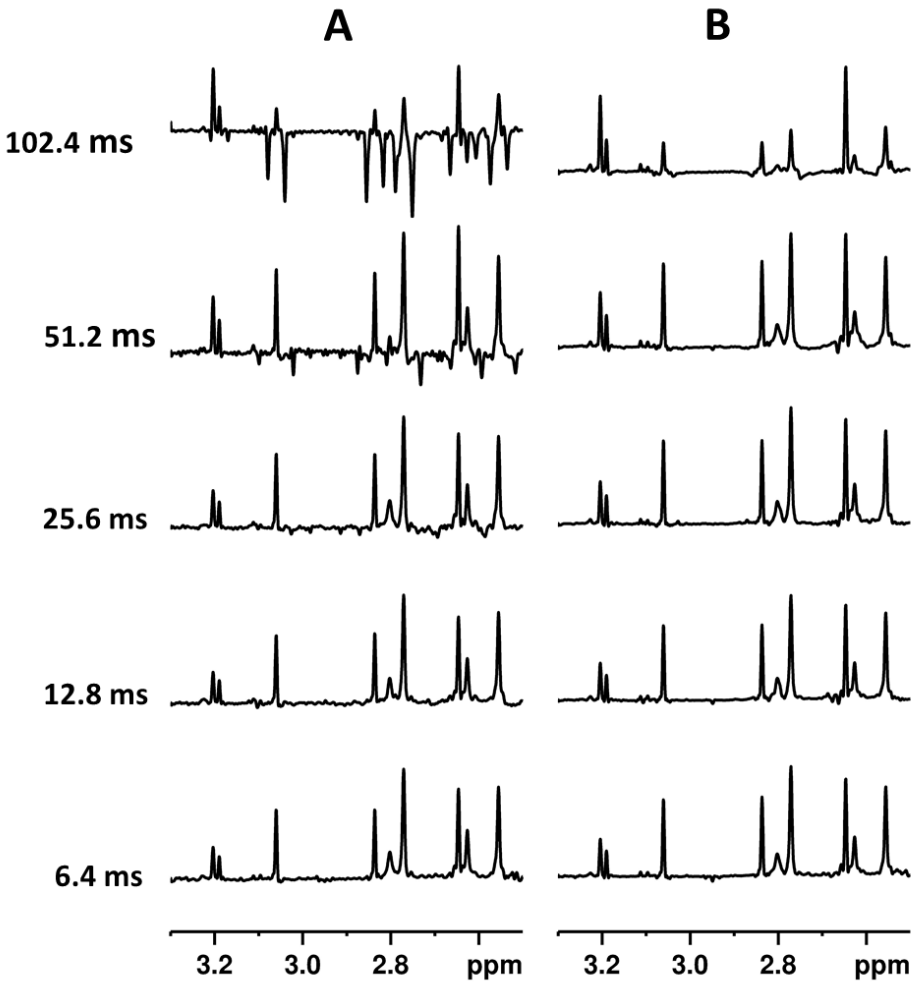
1. Hall, R., Beale, M., Fiehn, O., Hardy, N., Sumner, L., Bino, R. Plant metabolomics: the missing link in functional genomics strategies. *The Plant Cell*. **14** (7), 1437–1440, doi: 10.1105/tpc.140720 (2002).
2. Fiehn, O. Metabolomics-the link between genotypes and phenotypes. *Plant Molecular Biology*. **48** (1–2), 155–171 (2002).
3. Schauer, N., Fernie, A.R. Plant metabolomics: towards biological function and mechanism.

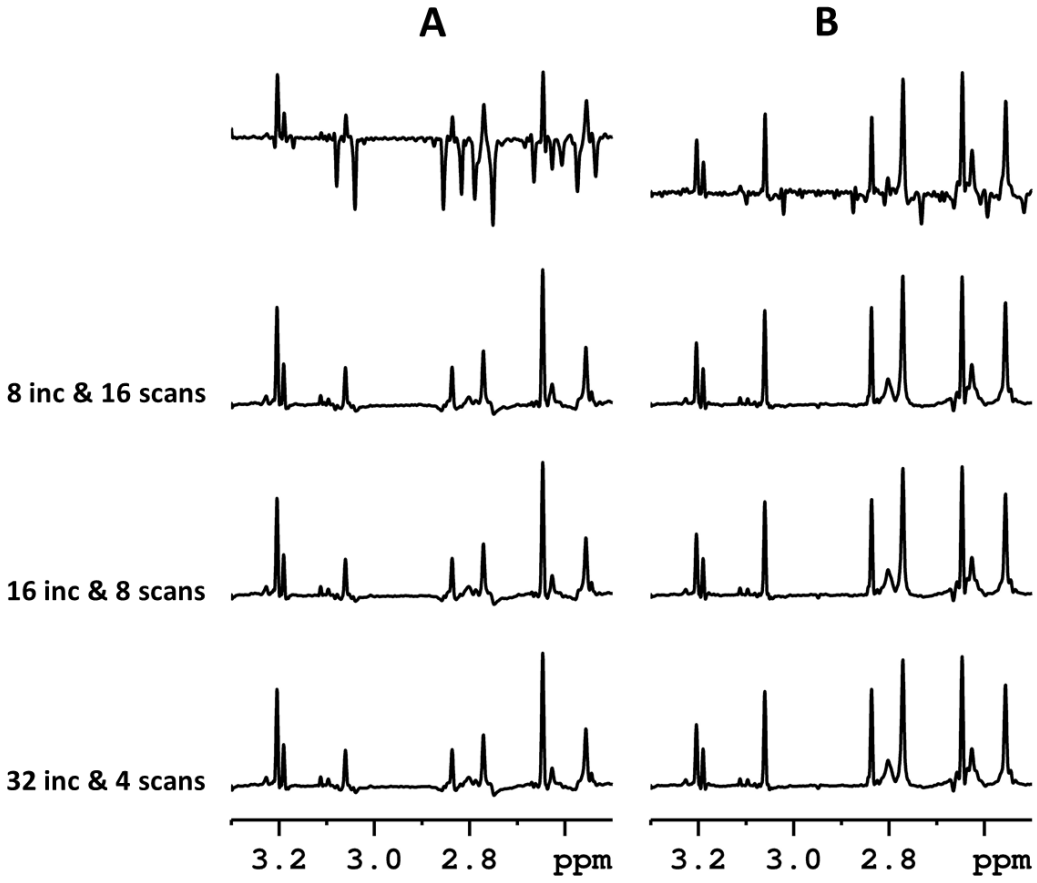
Trends in Plant Science. **11** (10), 508–516, doi: 10.1016/j.tplants.2006.08.007 (2006).
4. Kim, H.K., Choi, Y.H., Verpoorte, R. NMR-based plant metabolomics: where do we stand, where do we go? *Trends in Biotechnology*. **29** (6), 267–275, doi: 10.1016/j.tibtech.2011.02.001 (2011).
5. Kumar, R., Bohra, A., Pandey, A.K., Pandey, M.K., Kumar, A. Metabolomics for Plant Improvement: Status and Prospects. *Frontiers in Plant Science*. **8**, doi: 10.3389/fpls.2017.01302 (2017).
6. Dumez, J.-N. *et al.* Hyperpolarized NMR of plant and cancer cell extracts at natural abundance. *Analyst*. **140** (17), 5860–5863, doi: 10.1039/C5AN01203A (2015).
7. Emwas, A.-H. *et al.* NMR Spectroscopy for Metabolomics Research. *Metabolites*. **9** (7), doi: 10.3390/metabo9070123 (2019).
8. Zangger, K. Pure shift NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy*. **86–87**, 1–20, doi: 10.1016/j.pnmrs.2015.02.002 (2015).
9. Foroozandeh, M., Morris, G.A., Nilsson, M. PSYCHE Pure Shift NMR Spectroscopy. *Chemistry – A European Journal*. **24** (53), 13988–14000, doi: https://doi.org/10.1002/chem.201800524 (2018).
10. Castañar, L. Pure shift NMR: Past, present, and future. *Magnetic Resonance in Chemistry*. **56** (10), 874–875, doi: https://doi.org/10.1002/mrc.4758 (2018).
11. Lopez, J.M., Cabrera, R., Maruenda, H. Ultra-Clean Pure Shift 1 H-NMR applied to metabolomics profiling. *Scientific Reports*. **9** (1), 1–8, doi: 10.1038/s41598-019-43374-5 (2019).
12. Marcó, N., Gil, R.R., Parella, T. Isotropic/Anisotropic NMR Editing by Resolution-Enhanced NMR Spectroscopy. *Chemphyschem: A European Journal of Chemical Physics and Physical Chemistry*. **19** (9), 1024–1029, doi: 10.1002/cphc.201800094 (2018).
13. Kaltschnee, L. *et al.* Extraction of distance restraints from pure shift NOE experiments. *Journal of Magnetic Resonance*. **271**, 99–109, doi: 10.1016/j.jmr.2016.08.007 (2016).
14. Sinnaeve, D. *et al.* Improved Isotopic Profiling by Pure Shift Heteronuclear 2D J-Resolved NMR Spectroscopy. *Analytical Chemistry*. **90** (6), 4025–4031, doi: 10.1021/acs.analchem.7b05206 (2018).
15. Timári, I. *et al.* Real-Time Pure Shift HSQC NMR for Untargeted Metabolomics. *Analytical Chemistry*. **91** (3), 2304–2311, doi: 10.1021/acs.analchem.8b04928 (2019).
16. Zhao, Q. *et al.* Combination of pure shift NMR and chemical shift selective filters for analysis of Fischer-Tropsch waste-water. *Analytica Chimica Acta*. **1110**, 131–140, doi: 10.1016/j.aca.2020.03.014 (2020).
17. Zhao, Q. *et al.* Pure Shift NMR: Application of 1D PSYCHE and 1D TOCSY-PSYCHE Techniques for Directly Analyzing the Mixtures from Biomass-Derived Platform Compound Hydrogenation/Hydrogenolysis. *ACS Sustainable Chemistry & Engineering*. **9** (6), 2456–2464, doi: 10.1021/acssuschemeng.0c06882 (2021).
18. Foroozandeh, M. *et al.* Ultrahigh-Resolution Diffusion-Ordered Spectroscopy. *Angewandte Chemie International Edition*. **55** (50), 15579–15582, doi: https://doi.org/10.1002/anie.201609676 (2016).
19. Castañar, L., Pérez-Trujillo, M., Nolis, P., Monteagudo, E., Virgili, A., Parella, T. Enantiodifferentiation through Frequency-Selective Pure-Shift 1H Nuclear Magnetic Resonance Spectroscopy. *ChemPhysChem*. **15** (5), 854–857, doi: 10.1002/cphc.201301130 (2014).
20. Lopez, J.M., Sánchez, L.F., Nakamatsu, J., Maruenda, H. Study of the Acetylation Pattern of

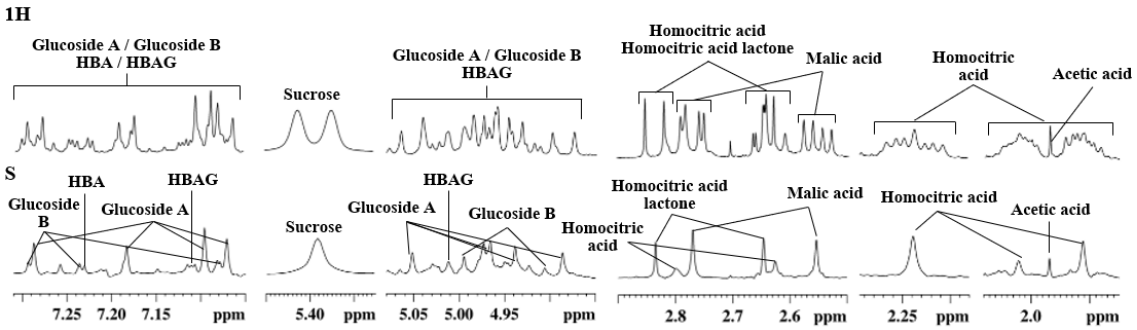
Chitosan by Pure Shift NMR. *Analytical Chemistry*. doi: 10.1021/acs.analchem.0c01638 (2020).
21. Foroozandeh, M., Adams, R.W., Meharry, N.J., Jeannerat, D., Nilsson, M., Morris, G.A. Ultrahigh-Resolution NMR Spectroscopy. *Angewandte Chemie International Edition*. **53** (27), 6990–6992, doi: 10.1002/anie.201404111 (2014).
22. Foroozandeh, M., Adams, R.W., Kiraly, P., Nilsson, M., Morris, G.A. Measuring couplings in crowded NMR spectra: pure shift NMR with multiplet analysis. *Chemical Communications*. **51** (84), 15410–15413, doi: 10.1039/C5CC06293D (2015).
23. Moutzouri, P. *et al.* Ultraclean pure shift NMR. *Chemical Communications*. **53** (73), 10188–10191, doi: 10.1039/C7CC04423B (2017).
24. Santacruz, L., Hurtado, D.X., Doohan, R., Thomas, O.P., Puyana, M., Tello, E. Metabolomic study of soft corals from the Colombian Caribbean: PSYCHE and 1 H-NMR comparative analysis. *Scientific Reports*. **10** (1), 5417, doi: 10.1038/s41598-020-62413-0 (2020).
25. Stark, P., Zab, C., Porzel, A., Franke, K., Rizzo, P., Wessjohann, L.A. PSYCHE—A Valuable Experiment in Plant NMR-Metabolomics. *Molecules*. **25** (21), 5125, doi: 10.3390/molecules25215125 (2020).
26. Kakita, V.M.R., Rachineni, K., Hosur, R.V. Ultraclean Pure Shift NMR Spectroscopy with Adiabatic Composite Refocusing Pulses: Application to Metabolite Samples. *ChemistrySelect*. **4** (34), 9893–9896, doi: https://doi.org/10.1002/slct.201902238 (2019).
27. Bo, Y. *et al.* High-resolution pure shift NMR spectroscopy offers better metabolite discrimination in food quality analysis. *Food Research International*. **125**, 108574, doi: 10.1016/j.foodres.2019.108574 (2019).
28. Watermann, S., Schmitt, C., Schneider, T., Hackl, T. Comparison of Regular, Pure Shift, and Fast 2D NMR Experiments for Determination of the Geographical Origin of Walnuts. *Metabolites*. **11** (1), 39, doi: 10.3390/metabo11010039 (2021).
29. Leyva-Zegarra, V. *et al.* NMR-based leaf metabolic profiling of V. planifolia and three endemic Vanilla species from the Peruvian Amazon. *Food Chemistry*. 129365, doi: 10.1016/j.foodchem.2021.129365 (2021).
30. Toubiana, D. *et al.* Morphological and metabolic profiling of a tropical-adapted potato association panel subjected to water recovery treatment reveals new insights into plant vigor. *The Plant Journal*. **103** (6), 2193–2210, doi: https://doi.org/10.1111/tpj.14892 (2020).
31. Maruenda, H., Cabrera, R., Cañari-Chumpitaz, C., Lopez, J.M., Toubiana, D. NMR-based metabolic study of fruits of Physalis peruviana L. grown in eight different Peruvian ecosystems. *Food Chemistry*. **262**, 94–101, doi: 10.1016/j.foodchem.2018.04.032 (2018).
32. Cloarec, O. *et al.* Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Analytical Chemistry*. **77** (5), 1282–1289, doi: 10.1021/ac048630x (2005).
33. Steuer, R. Review: on the analysis and interpretation of correlations in metabolomic data. *Briefings in Bioinformatics*. **7** (2), 151–158, doi: 10.1093/bib/bbl009 (2006).
34. Trygg, J., Holmes, E., Lundstedt, T. Chemometrics in Metabonomics. *Journal of Proteome Research*. **6** (2), 469–479, doi: 10.1021/pr060594q (2007).
35. Worley, B., Powers, R. Multivariate Analysis in Metabolomics. *Current Metabolomics*. **1** (1), 92–107, doi: 10.2174/2213235X11301010092 (2013).

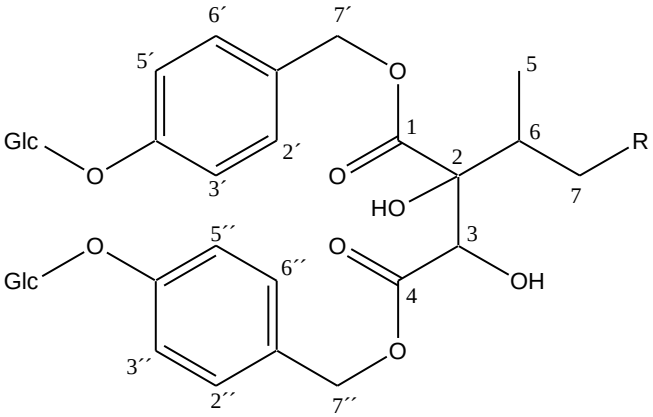
Figure




A
B
102.4 ms
51.2 ms
25.6 ms
12.8 ms
6.4 ms
3.2
3.0
2.8
ppm
3.2
3.0
2.8
ppm

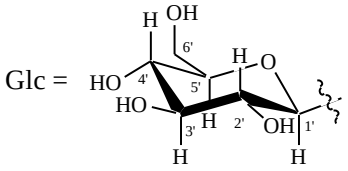

A
B
8 inc & 16 scans
16 inc & 8 scans
32 inc & 4 scans
3.2
3.0
2.8
ppm
3.2
3.0
2.8
ppm

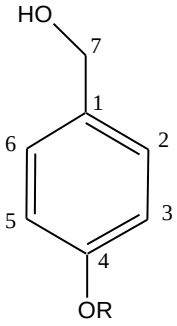

1H
Glucoside A / Glucoside B
HBA / HBAG
Sucrose
Glucoside A / Glucoside B
HBAG
Homocitric acid
Homocitric acid lactone
Malic acid
Homocitric
acid
Acetic acid
S
HBA
HBAG
Glucoside
B
Glucoside A
Sucrose
HBAG
Glucoside A
Glucoside B
Homocitric
acid
Homocitric acid
lactone
Malic acid
Homocitric acid
Acetic acid
7.25 7.20 7.15 ppm
5.40 ppm
5.05 5.00 4.95 ppm
2.8 2.7 2.6 ppm
2.25 ppm
2.0 ppm



R=H, Glucoside A

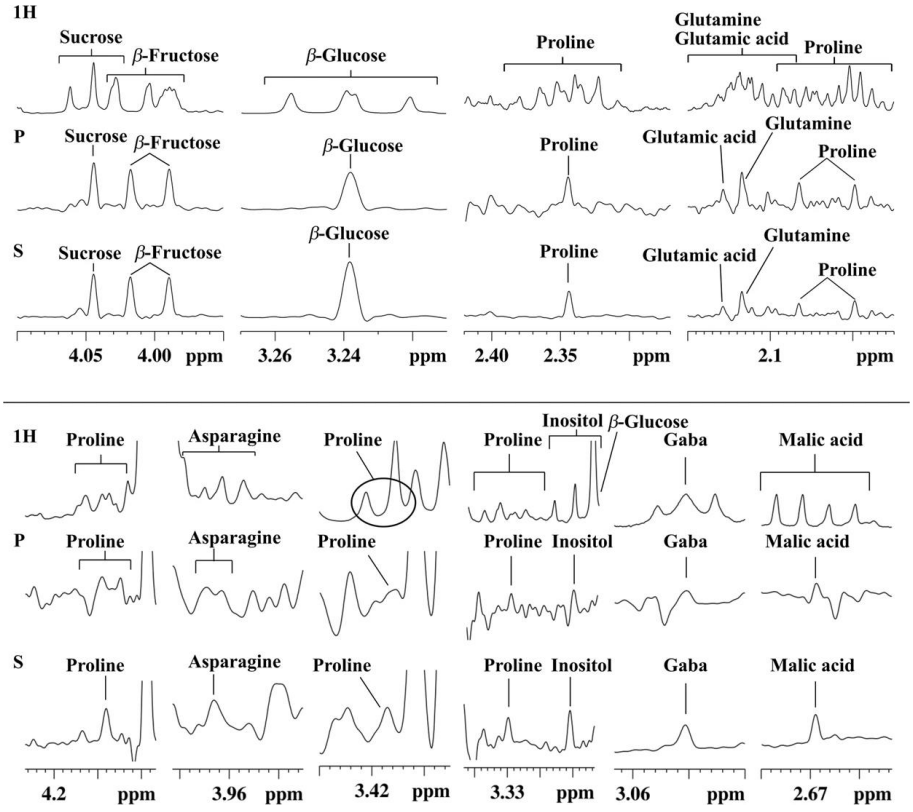
R=CH_3, Glucoside B



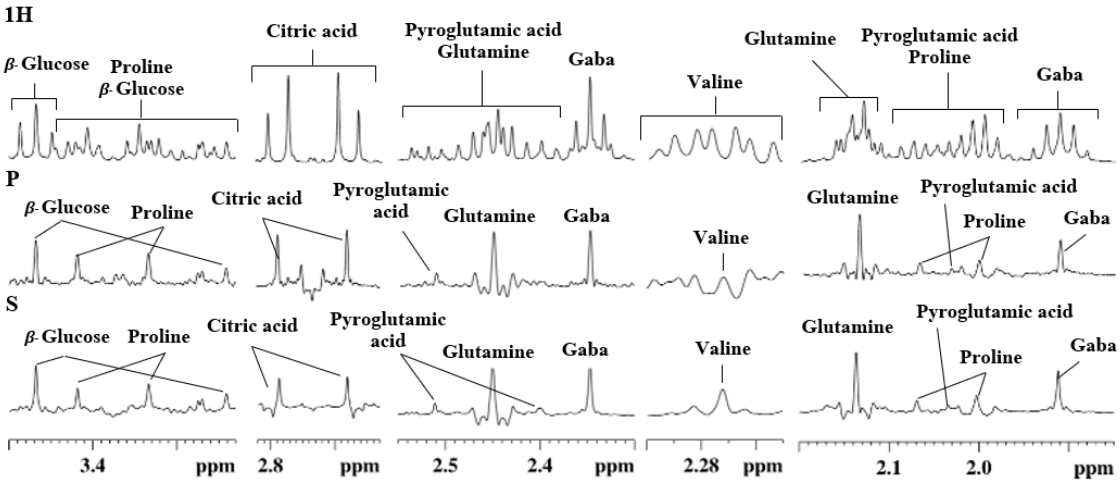


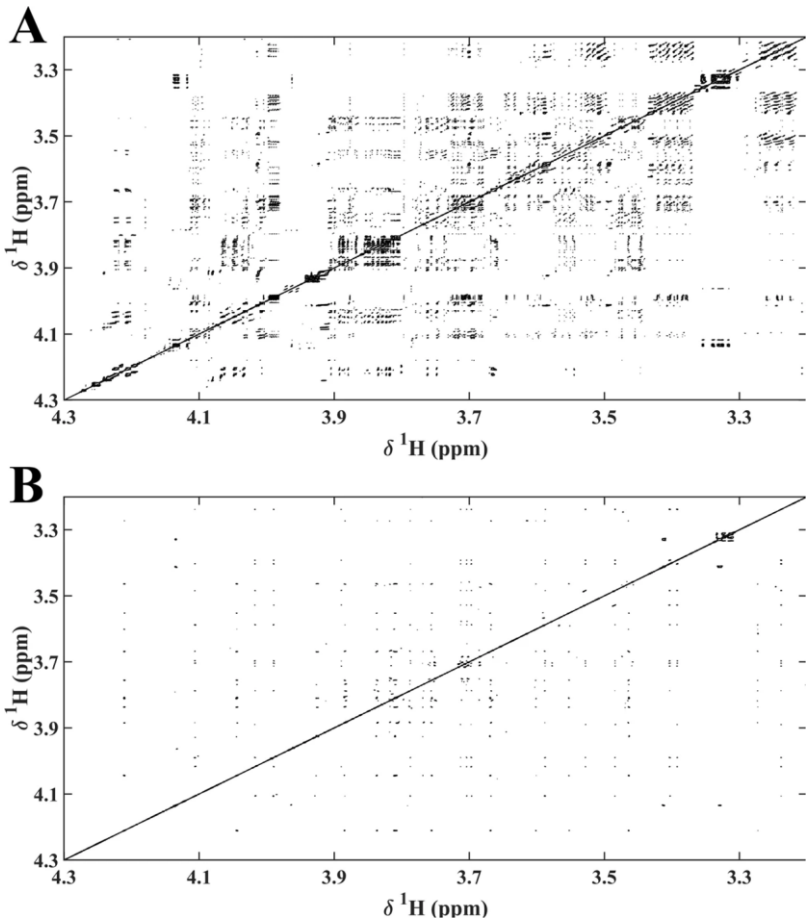
R=H, 4-hydroxybenzyl alcohol (4-HBA)

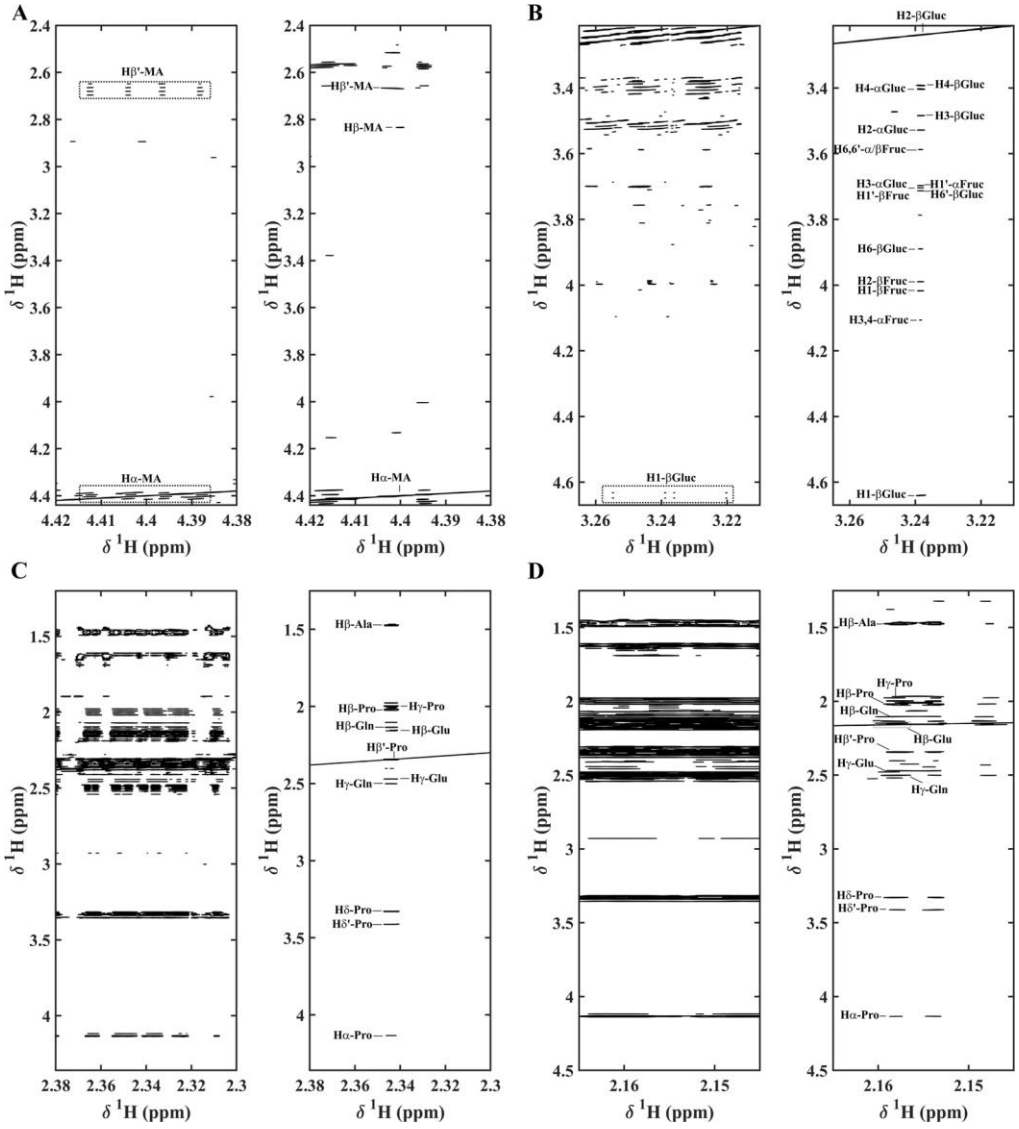
R= Glc, 4-HBA glucoside

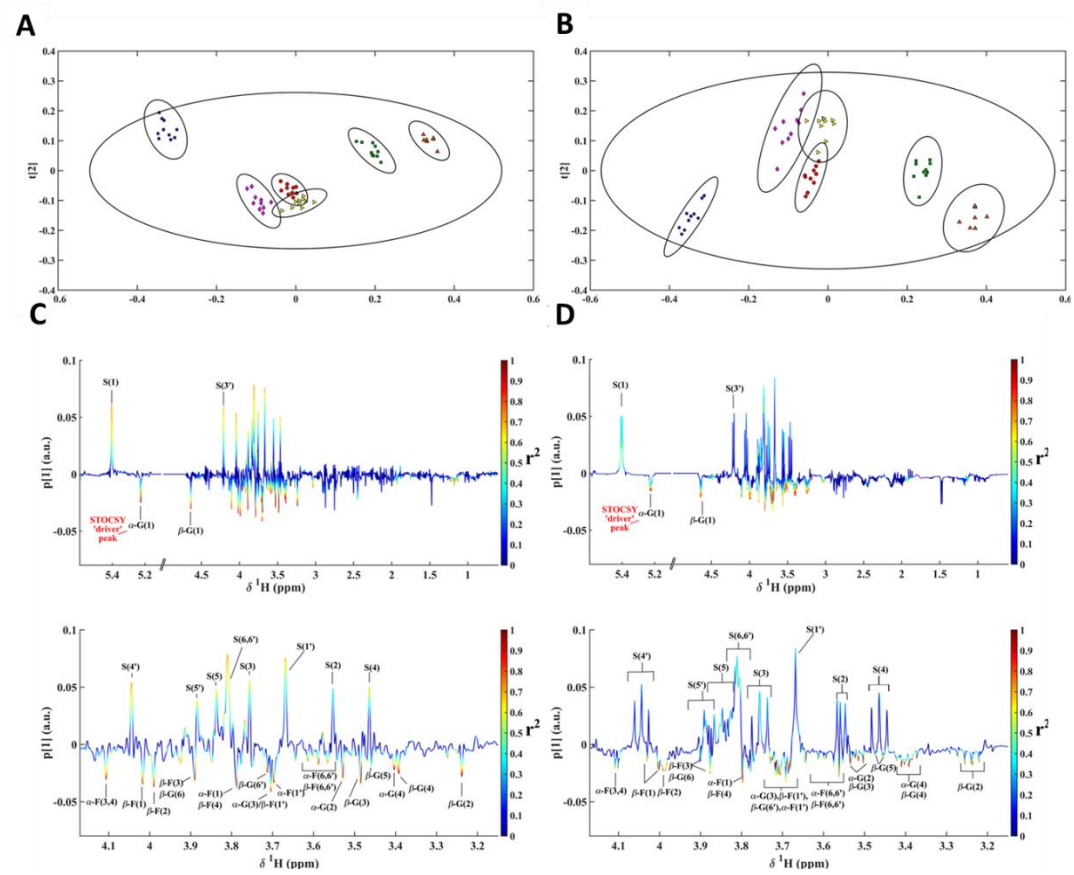

1H
Sucrose
β-Fructose
β-Glucose
Proline
Glutamine
Glutamic acid
Proline
P
Sucrose
β-Fructose
β-Glucose
Proline
Glutamic acid
Glutamine
Proline
S
Sucrose
β-Fructose
β-Glucose
Proline
Glutamic acid
Glutamine
Proline
4.05
4.00
ppm
3.26
3.24
ppm
2.40
2.35
ppm
2.1
ppm

1H
Proline
Asparagine
Proline
Inositol
β-Glucose
Proline
Gaba
Malic acid
P
Proline
Asparagine
Proline
Proline
Inositol
Gaba
Malic acid
S
Proline
Asparagine
Proline
Proline
Inositol
Gaba
Malic acid
4.2
ppm
3.96
ppm
3.42
ppm
3.33
ppm
3.06
ppm
2.67
ppm


1H
β-Glucose
Proline
β-Glucose
Citric acid
Pyroglutamic acid
Glutamine
Gaba
Valine
Glutamine
Pyroglutamic acid
Proline
Gaba
P
β-Glucose
Proline
Citric acid
Pyroglutamic acid
Glutamine
Gaba
Valine
Glutamine
Pyroglutamic acid
Proline
Gaba
S
β-Glucose
Proline
Citric acid
Pyroglutamic acid
Glutamine
Gaba
Valine
Glutamine
Pyroglutamic acid
Proline
Gaba
3.4
ppm
2.8
ppm
2.5
2.4
ppm
2.28
ppm
2.1
2.0
ppm


A
3.3
3.5
3.7
3.9
4.1
4.3
δ 1H (ppm)
4.3
4.1
3.9
3.7
3.5
3.3
δ 1H (ppm)
B
3.3
3.5
3.7
3.9
4.1
4.3
δ 1H (ppm)
4.3
4.1
3.9
3.7
3.5
3.3
δ 1H (ppm)


A
Hβ'-MA
Hα-MA
Hβ-MA
B
H1-βGluc
H2-βGluc
H4-αGluc
H4-βGluc
H3-βGluc
H2-αGluc
H6,6'-α/βFruc
H3-αGluc
H1'-αFruc
H1'-βFruc
H6'-βGluc
H6-βGluc
H2-βFruc
H1-βFruc
H3,4-αFruc
H1-βGluc
C
Hβ-Ala
Hβ-Pro
Hγ-Pro
Hβ-Gln
Hβ-Glu
Hβ'-Pro
Hγ-Gln
Hγ-Glu
Hδ-Pro
Hδ'-Pro
Hα-Pro
D
Hβ-Ala
Hγ-Pro
Hβ-Pro
Hβ-Gln
Hβ'-Pro
Hβ-Glu
Hγ-Glu
Hγ-Gln
Hδ-Pro
Hδ'-Pro
Hα-Pro
δ 1H (ppm)


A
B
C
D
t[2]
p[1] (a.u.)
δ ¹H (ppm)
r²
S(1)
S(3')
STOCSY 'driver' peak
α-G(1)
β-G(1)
S(4')
S(5')
S(5)
S(6,6')
S(3)
S(1')
S(2)
S(4)
α-F(3,4)
β-F(1)
β-F(2)
β-F(3)
β-G(6)
α-F(1)
β-F(4)
β-G(6')
α-G(3)/β-F(1')
α-F(1')
α-F(6,6')
β-F(6,6')
α-G(2)
β-G(3)
β-G(5)
α-G(4)
β-G(4)
β-G(2)
α-G(3),β-F(1'),
β-G(6'),α-F(1')
α-F(6,6')
β-F(6,6')
α-G(2)
β-G(3)
β-G(5')
α-G(4)
β-G(4)

Click here to access/download

Table of Materials

Materials.xls

June the 14th, 2021

No JoVE62719
Pure Shift Nuclear Magnetic Resonance: a new tool for plant metabolomics.,

Dear Editor:

We would like to thank you and the reviewers for their time in revising our manuscript. In the new version here attached all the questions raised by the reviewers have been addressed.

The resubmission includes the following files as requested:

(1) a cover letter indicating ITEMIZED NUMBERED RESPONSE to each of the reviewer's comments. Comments by the reviewers are provided in black normal font, our responses are highlighted in blue font, and the in-text citations of the updated version of the manuscript are shown in red. Black bold font is used to highlight original text in submitted manuscript.
(2) a clean unmarked copy of the revised manuscript.
(3) All figures and tables are included separately.
(4) The Supplementary Information has been updated.

We appreciate the support received to improve the quality of this document.

On behalf of the authors,

Dr. Juan Manuel Lopez

RESPONSE TO REVIEWER COMMENTS:

Reviewer: 1

Manuscript Summary:
This manuscript aims to provide a practical guide how to set up PSYCHE and SAPPHIRE-PSYCHE for metabolomics studies. It provides some examples to illusrate the advantages of the increased resolution of and of the chunking artefact suppression for the analysis. It also refers to previous papers published by the authors. I believe this paper can indeed be useful for analytical chemists who are not necessarily NMR experts, but who wish to apply pure shift methods for their metabolomics studies.

I believe the paper can be published without major revisisions. I have a number of minor remarks that the authors should consider, mostly comments on the way things are phrased that may create confusion.

Copy-editing: The writing is mostly fine, but I did encounter a few typographical errors here and there.

We thank the reviewer for taking the time of detecting several typo errors. We really appreciate it.

Major Concerns:
None.

Minor Concerns:

1- line 75 : authors mean a "crowded spectrum"

Correction made.

Line 73-75: First, all the 1H NMR signals detected in the sample are distributed in a small window corresponding to the proton chemical shift window, which results in crowded spectra.

2- lines 98-99 : Authors mean "sideband artifacts" and not "size band artifacts". Also, I propose to rephrase this sentence, as it subtly confuses two concepts: it is not the interferogram acquisition that leads to J-modulation artefacts, but the fact that data is acquired as chunks. If the interferogram would be performed with the acquisition of single data points per increment, there would be no such sidebands (but of course that results in an unacceptably long experiment). My suggestion: "However, as PSYCHE is a 2D interferogram experiment where chunks of time domain data are acquired, it suffers from periodic sideband artifacts that result from J-coupling modulation distortions at the edges of these chunks."

We agree with the reviewer and the suggested corrections were made.

Line 94-99: In 2014, Foroozandeh et al. published a new Pure Shift experiment, PSYCHE (Pure Shift Yielded by Chirp Excitation), based on anti-z-COSY pulse sequence which yielded excellent homonuclear decoupling and improved sensitivity values.[21] However, as PSYCHE is a 2D interferogram experiment where chunks of time domain data are acquired, it suffers from periodic sideband artifacts that result from J-coupling modulation distortions at the edges of the chunk.

3-line 101: "without sensitivity penalty". Strictly speaking, there is a sensitivity penalty, since (1) the additional delays in the pulse sequence causes some additional T2 relaxation and (2) the averaging of the differently J-modulated time domain data. However, these losses are not so high and the gain in spectral purity by far outweighs this disadvantage. It would be worthwhile to modify this statement accordingly.
The same remark for the sentence at

We agree with the reviewer and we have made the change in the manuscript

Line 106-109: In 2019, we demonstrated for the first time11 that the SAPPHIRE-PSYCHE Pure Shift method, which removes artifacts with almost no sensitivity penalty[23]- could be employed for the analysis of complex biological mixtures, such as extracts of the fruits of Physalis peruviana, commonly known as Cape gooseberries.[11]

4- line 203: the chirp pulse flip angle needs to be "small", not "short" (using the word "short" would suggest changing the PSYCHE pulse length, which is not done here).

We agree with the reviewer and we have made the change in the manuscript

Line 205-207: Note: The PSYCHE experiment is based on an anti-z-COSY scheme, consequently the CHIRP pulse flip angle needs to be small to avoid recoupling artifacts (Figure 3).

5-lines 206 and 246 : the authors talk about "sensibility", but they mean "sensitivity".

We agree with the reviewer and we apologize. The corrections have been made.

Line 210-211: A good compromise between sensitivity and low recoupling artifacts is to set CNST61 = 20°.[19, 22]

Line 249-251: Note: As in the regular PSYCHE experiment, the CHIRP pulse flip angle needs to be short to avoid recoupling artifacts. CNST20 = 20° is a good compromise between sensitivity and low recoupling artifacts.[21, 23, 25]

6- line 211: "SPNAN" should be "SPNAM"

We agree with the reviewer and we have made the change in the manuscript

Line 215: Choose the Crp_psyche.20 (SPNAM 37) shape pulse for the PSYCHE element (Figure 2).

7- lines 213 and 252 : "small gradient" : the correct way to refer to this is a "weak magnetic field gradient".

We agree with the reviewer and we have made the change in the manuscript

Line 218-220: Note: A weak magnetic field gradient is applied during the PSYCHE element, normally, between 1% to 4% of the maximum strength of the gradient, depending on the probe.

Line 258-261: Note: A weak magnetic field gradient is applied during the PSYCHE element, normally, between 1% to 4% of the maximum strength of the gradient, depending on the probe.

8- line 226: the authors might want to mention that L31 should best be set to a power of 2, so that the final 1D has a power of 2 number of time domain points, as in regular 1D spectra (this is not a hard requirement, but gives maximum information content in the real part of the spectrum).

We agree with the reviewer and we have added the suggestion into the manuscript

Line 200-202: Note: L31 is the number of complex digital points acquired in each Pure Shift block, best to be set to a power of 2.[21]

9- lines 264-266: longer data chunks are indeed allowed with SAPPHIRE, but this will mean more averaging is needed to remove the now stronger sidebands. This discussion is in the original SAPPHIRE paper from Moutzouri et al. It would be good to mention this point, as it is here represented as if longer chunks come at no costs.

We agree with the reviewer and we have added the suggestion into the manuscript

Line 272-274: However, longer chunk data acquisition leads to higher J-coupling evolutions, which would require more J-coupling phase modulation increments to remove the stronger sidebands attained.[23]

10- line 336-337: same remark as before: "the artifacts generated during chunked data acquisition are a problem".

We agree with the reviewer and we have modified the paragraph

Line 342-246: The PSYCHE sequence is relatively easy to use and it has been successfully implemented in a wide range of applications.[9] To attain the decoupled spectrum, the pulse sequence acquires small chunks of FID refocussing the J-coupling at the middle of each block.

11- line 205 and line 434 (Figure 3): The authors should clearly state how the signal to noise was calculated, and show the 'noise' region that is used in the figure. If the noise region suffers from chunking artifacts, then I understand why the graphs in B and C do not show the same maximum. But then the signal to noise was not measured, but rather the signal to artifact ratio. For signal to noise, one should select a noise region with no signal and no artifact.

Indeed, the reviewer is correct, some artefact spread into the noise, we have stressed this point to avoid any confusion. The noise range used for the calculation was specified in figure 1 caption and show in supporting figure S12

Line 207-210: The absolute intensity increases with the excitation flip angle. The periodic artifacts are also enhanced, spreading into the spectrum and increasing the "noise" (Figure 1). The "noise" becomes a combination of standard noise and chuncking artifacts.

Caption figure 1: **The signal-to-noise ratio (S/N) (more precisely, signal to noise + artifact ratio) was calculated using the signal of maximum intensity against that of the noise, value calculated over a 2ppm range: from 7.75 ppm to 9.75 ppm (supporting figure S12).**

12- lines 635-636: same remark as before about the sensitivity penalty of SAPPHIRE. Given that the authors mention TD2*NS, what is probably meant here specifically is that there is in principle no cost in total experimental time given the same chunk lengths are used in SAPPHIRE-PSYCHE and PSYCHE. Experimental time and sensitivity are different concepts that should not be confused with one another.

We agree with the reviewer, to avoid any confusions, we have modified the paragraph.

Line 567-570: As periodic artifacts are strongly suppressed by SAPPHIRE, the block lengths could be longer; however, very long chunks strongly affect the intensity of the signals (Figures 2 and 3). **In SAPHIRE, J modulation increments contribute to spectrum sensitivity, therefore, the resulting decoupled FID total number of scans is equal to TD2 * NS (Figures 2 and 3).**[23]

Reviewer #2:
The paper presents an application of pure shift NMR as a new tool to the characterization of mixtures from plant metabolomics. The authors choose PSYCHE and SAPPHIRE-PSYCHE experiments to analyze Vanilla plant leaves, potato tubers (S. tuberosum) and Cape gooseberries (P. peruviana); and avoid signal overlaps. While this approach proves indeed effective and useful, the hyperfine structures cannot be observed.

1- I feel that alternate and much simpler techniques such as TSE-PSYCHE (doi.org/10.1039/C5CC06293D and doi.org/10.1016/j.aca.2020.03.014) would give a significant improvement in spectral purity and artifact suppression. I believe that the application of 1D PSYCHE pulse sequence has already been described in the literature. I am aware of this paper: doi.org/10.1021/acssuschemeng.0c06882.

We agree with the reviewer TSE-PSYCHE is an alternative method, not yet utilized in metabolomics. However, the purpose of this paper was to explain step by step the implementation of regular PSYCHE and SAPPHIRE-PSYCHE methodologies already described for metabolomics analysis (reference: Lopez et al. Ultra-Clean Pure Shift 1 H-NMR applied to metabolomics profiling. Scientific Reports. 9 (1), 1–8, doi: 10.1038/s41598-019-43374-5 (2019).
To comply with the reviewer, we included TSE-PSYCHE experiment and the reference in the manuscript (see reference 16, 17 and 22).

Line 101-104: **There are two methods to remove these artifacts - TSE-PHYCHE**[22] **and a more recent modification of the PSYCHE experiment called SAPPHIRE-PSYCHE (Sideband Averaging by Periodic PHase Incrementation of Residual J Evolution).**[23]

2- For the paper to be acceptable, the authors should give a normative name of PSYCHE pulse not "reset_psyche_1d", which cannot be founded from the Manchester NMR Methodology Group website.

Sorry for the confusion the PSYCHE experiment that we use is from bruker library as we mention in line XXXX

As SAPPHIRE–PSYCHE pulse sequence is not implemented in the Bruker library, we use the sequence downloaded from Manchester NMR group.
Line xxx

Reviewer #3:

Manuscript Summary:
The manuscript report on Protocol of the plant extract preparation in different food presented and their analysis using Pure Shift PSYCHE and SAPPHIRE-PSYCHE. Their comparison shown with the classical 1H-NMR. However, I would like to raise the following point that may contribute to improve the interest of this work.

Major Concerns:

1-Are Periodic size band artefact's not interfering with minor metabolites ?

Indeed, the reviewer is correct, sideband artifacts do interfere with the signals of minor metabolites. The purpose of this article is to show step by step how to implement SAPPHIRE –PSYCHE, method that removes sidebands artifacts that could overlap the signals of the minor compounds.

Minor Concerns:

2- Author should add discussion on quantitative analysis as author focused on qualitative aspect only.
Data about some references are not complete.

To comply with this request, we included a paragraph about quantification on the discussion.

Line 576-594: One limitation of both of these Pure Shift pulse sequences, PSYCHE and SAPPHIRE-PSYCHE, is quantification by absolute metabolite-internal standard integration. In regular 1H-NMR, the integrated intensity is directly proportional to the concentration of each metabolite. In PSYCHE, this is no longer the case, because a number of phenomena distort the signals and affect integration. For instance, the total integral value diminishes due to T2 relaxation during the pulse sequence spin selection. Also, the truncated J coupling evolution during the chunk acquisition which generates sidebands artifacts, disrupts the Lorentzian shape of the signal. Hence, the integral is now comprised by areas under the main peak and under all the sidebands, complicating signal integration.[8, 9, 21, 23] The frequency and magnitude of the sidebands are directly related to the chunk length but also to intrinsically molecular properties such as relaxation and the J coupling magnitude and multiplicity: higher J coupling magnitudes and higher multiplicities, lead to more distorted signals. In the case of SAPPHIRE, even though this NMR experiment efficiently removes the sideband artifacts, the signal intensities are compromised by the truncated J coupling evolution. The sum of each J modulated increment, generates an averaged decoupled FID where signal diminishment is directly related to the chunk length and the J coupling magnitude and multiplicity.[23] Moreover, CHIRP pulse flip angles generate recoupling artifact which also affects each signal differently, further complicating the quantification.[21] The magnitude of the effect of these

pulse sequences in quantitative analysis was assessed in our earlier Cape gooseberry study yielding errors of around 10% to 30%.[11]

3-Please add software used for PCA analysis.

To comply with this request, the software used for PCA analysis has been added to the manuscript. Caption figure 10:

Caption Figure 10: PLS-DA and STOCSY analysis was perform using MATLAB Version R2018a (Mathworks, Natick, MA).

Pure Shift Nuclear Magnetic Resonance: a new tool for plant metabolomics.

Juan M. Lopez[1], Vanessa Leyva[1], Helena Maruenda[1]
[1]Departamento de Ciencias – Química, CERMN, Pontificia Universidad Católica del Perú

Correspondence to: Juan M. Lopez at juan.lopez@pucp.edu.pe, Helena Maruenda at hmaruen@pucp.edu.pe

SUPPORTING FIGURES:

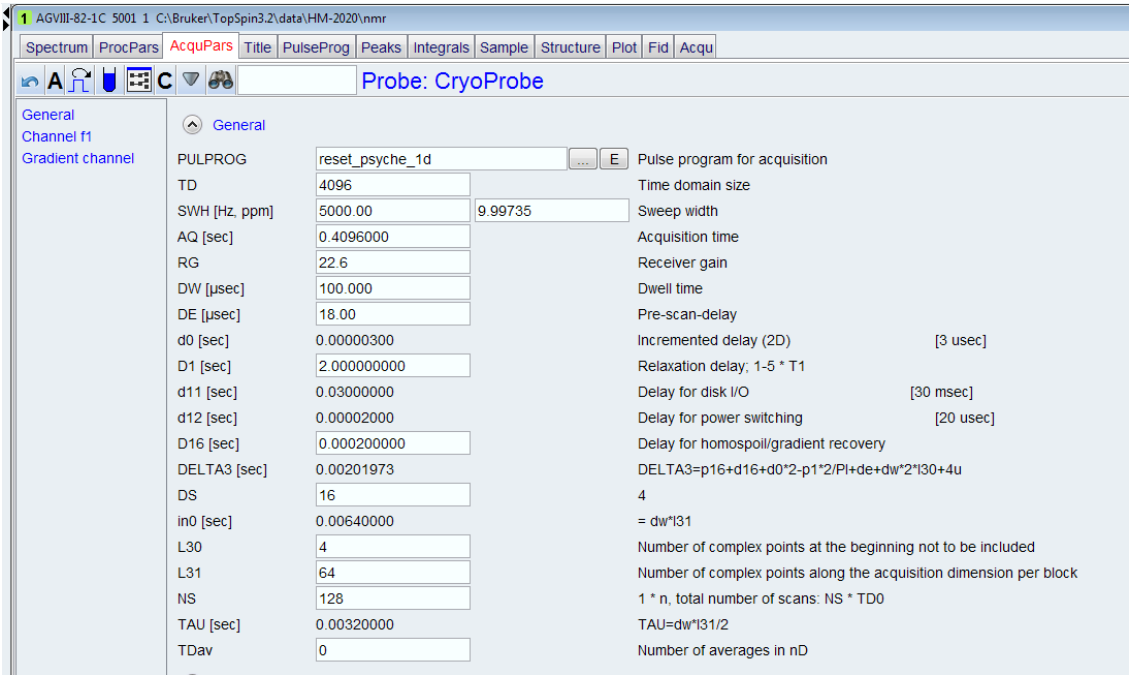


Figure S1. PSYCHE, standard acquisition parameters.

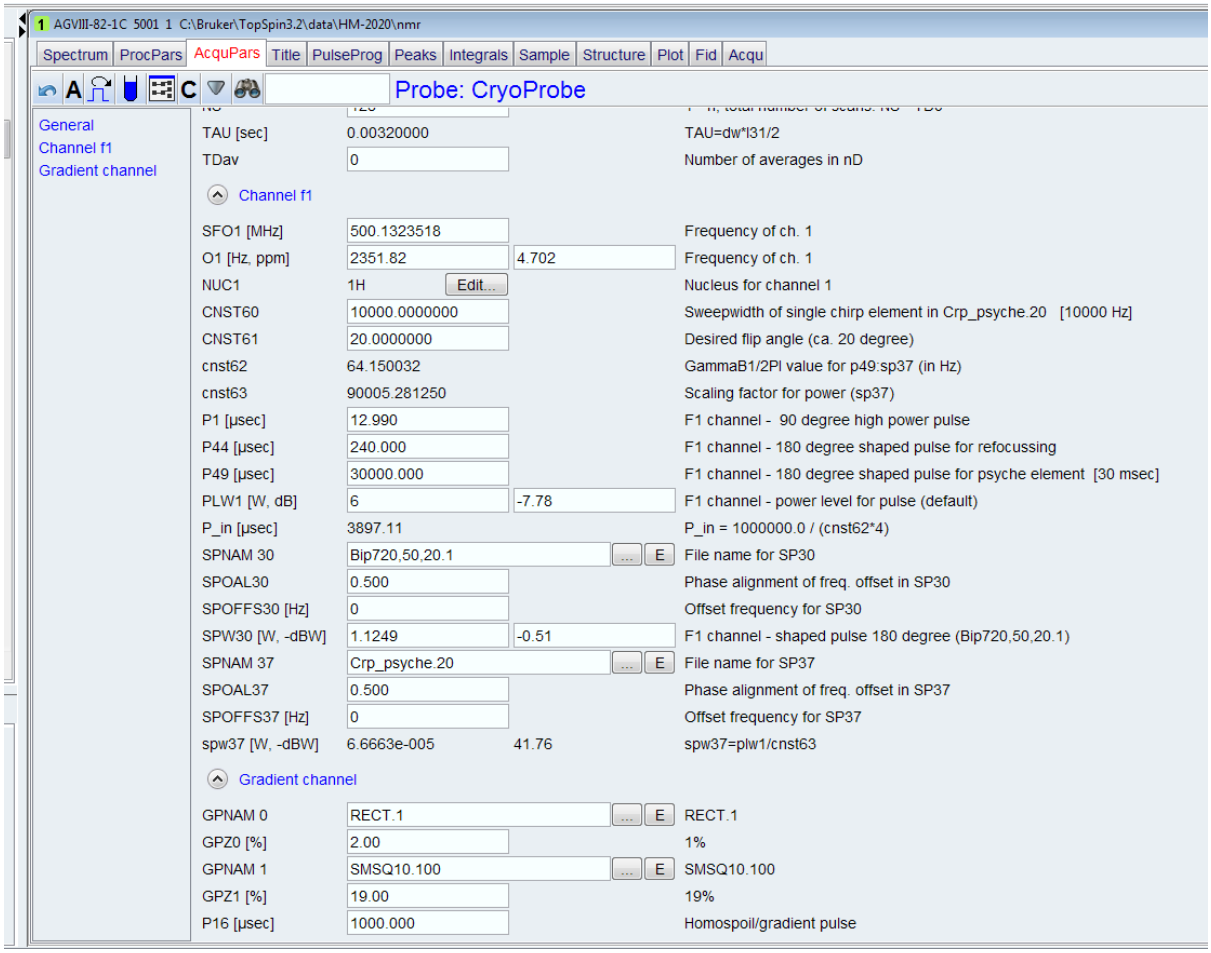


Figure S2. PSYCHE, standard acquisition parameters.

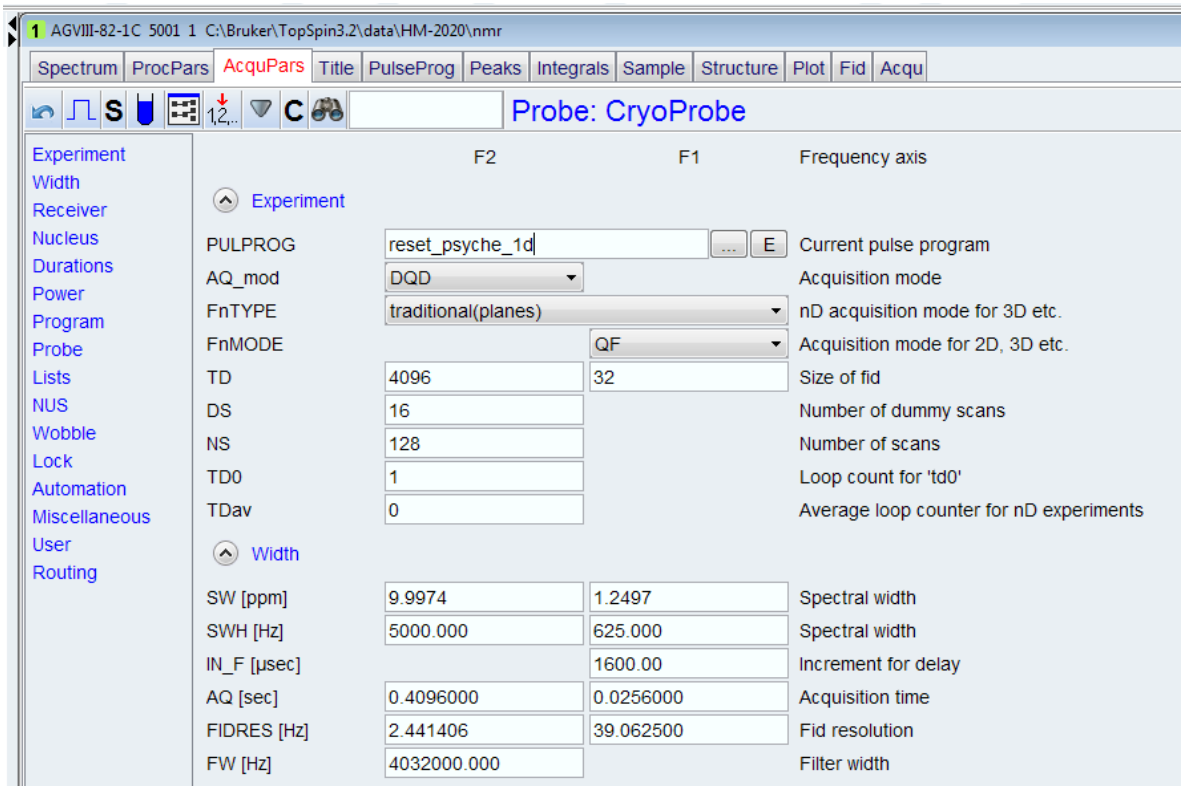


Figure S3. PSYCHE, standard acquisition parameters.

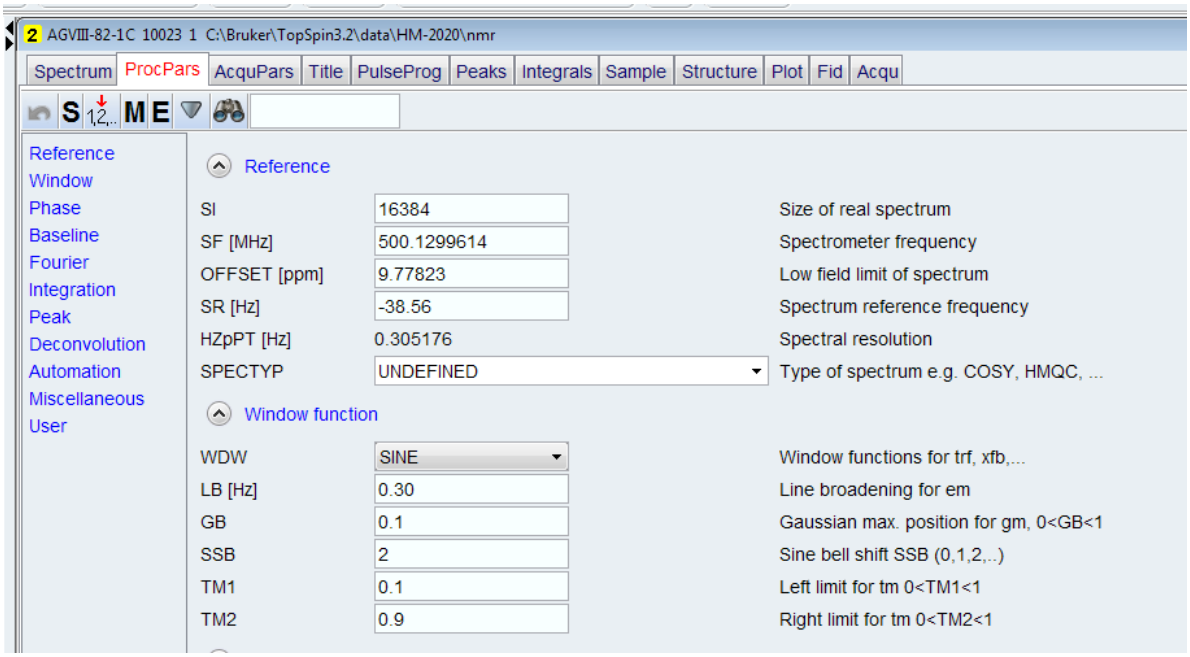


Figure S4. PSYCHE and SAPPHIRE-PSYCHE, standard processing parameters.

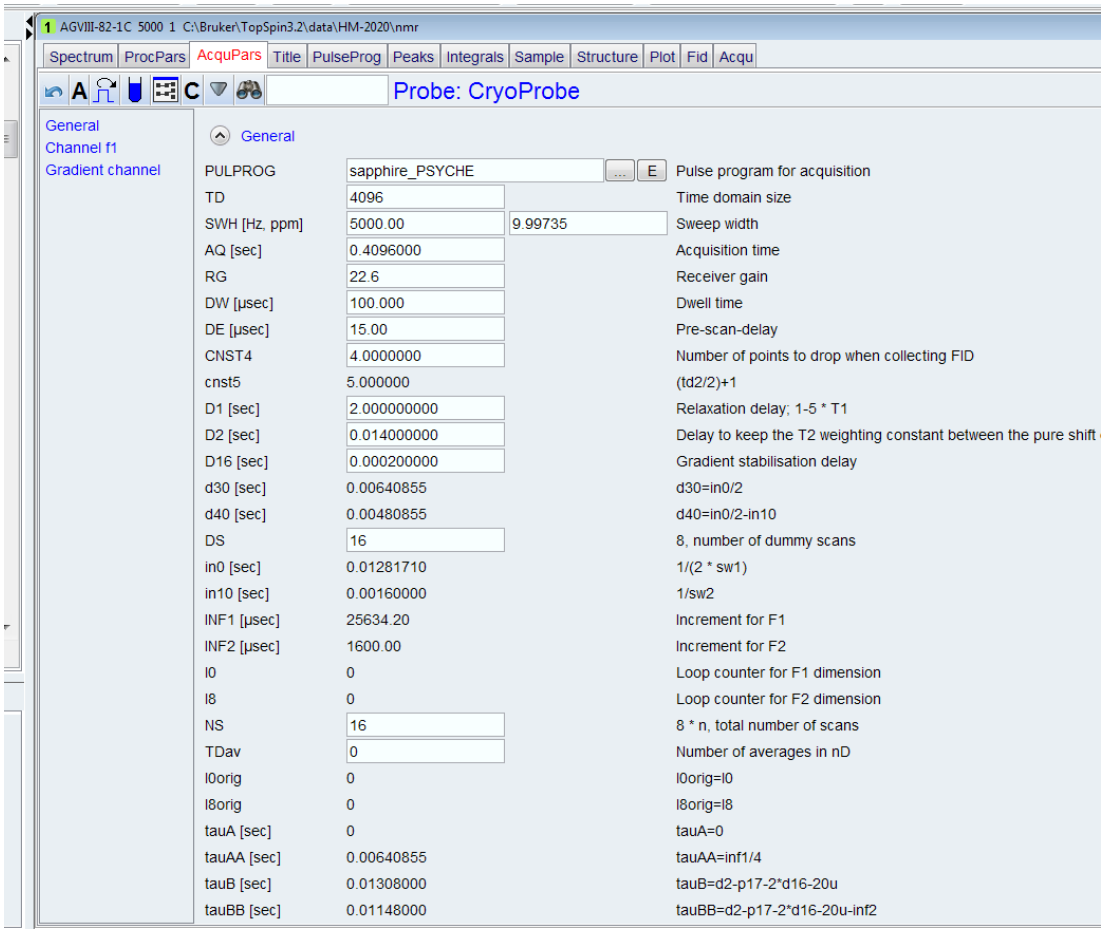


Figure S5. SAPPHIRE-PSYCHE, standard acquisition parameters.

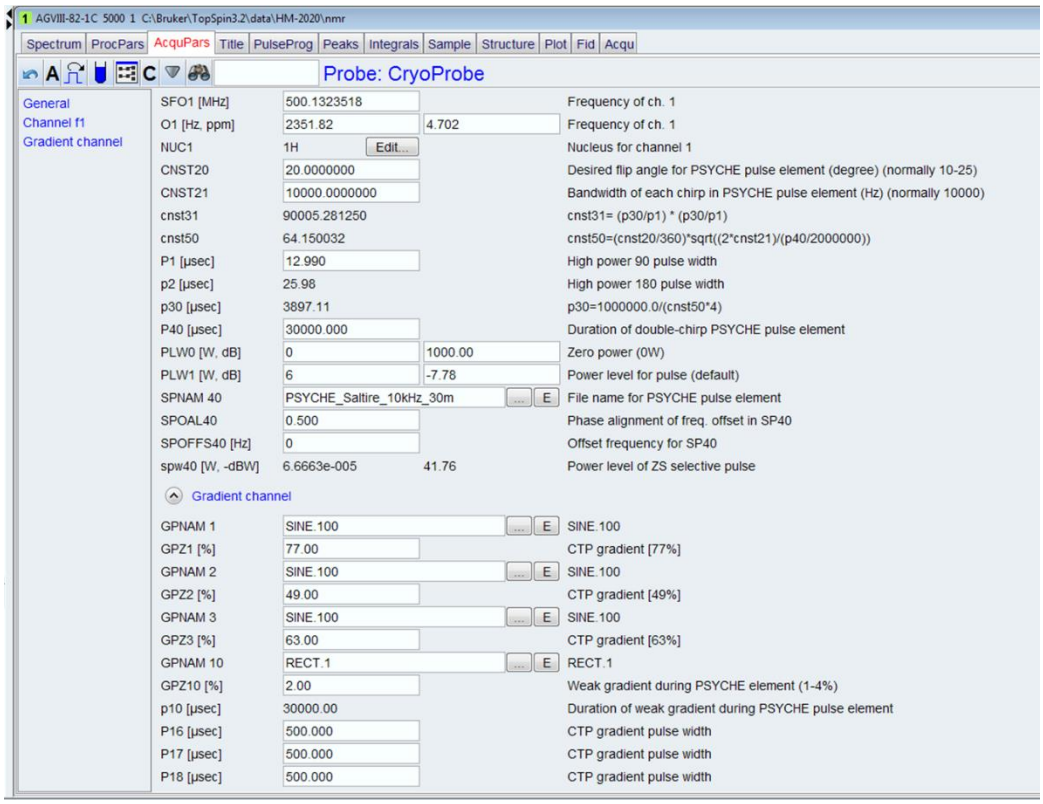


Figure S6. SAPPHIRE-PSYCHE, standard acquisition parameters.

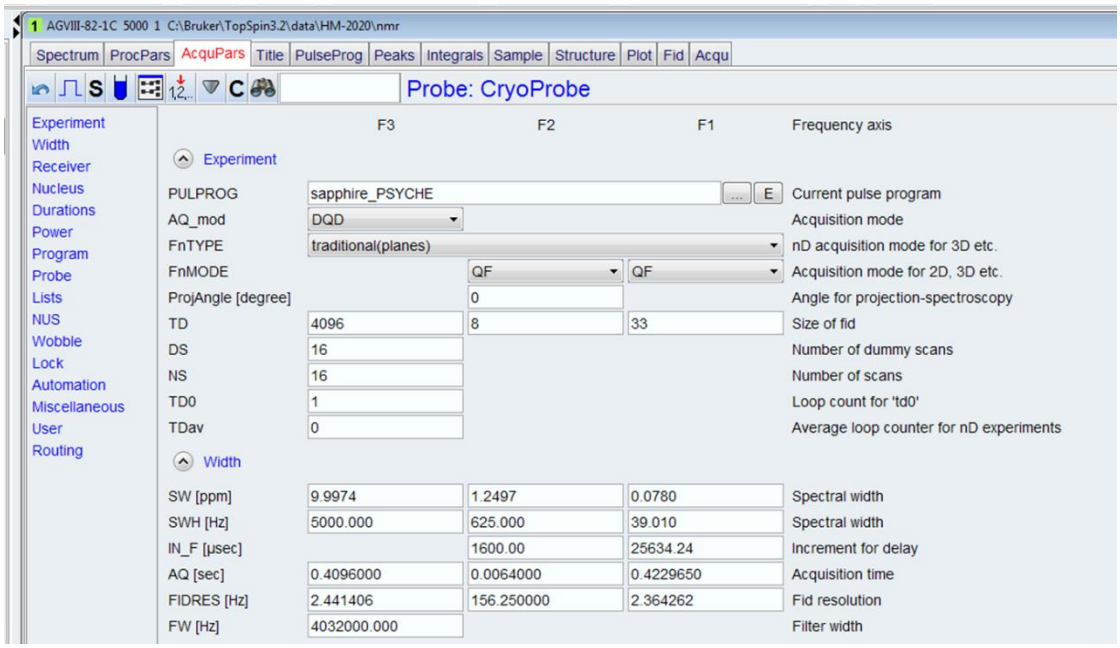


Figure S7. SAPPHIRE-PSYCHE, standard acquisition parameters.

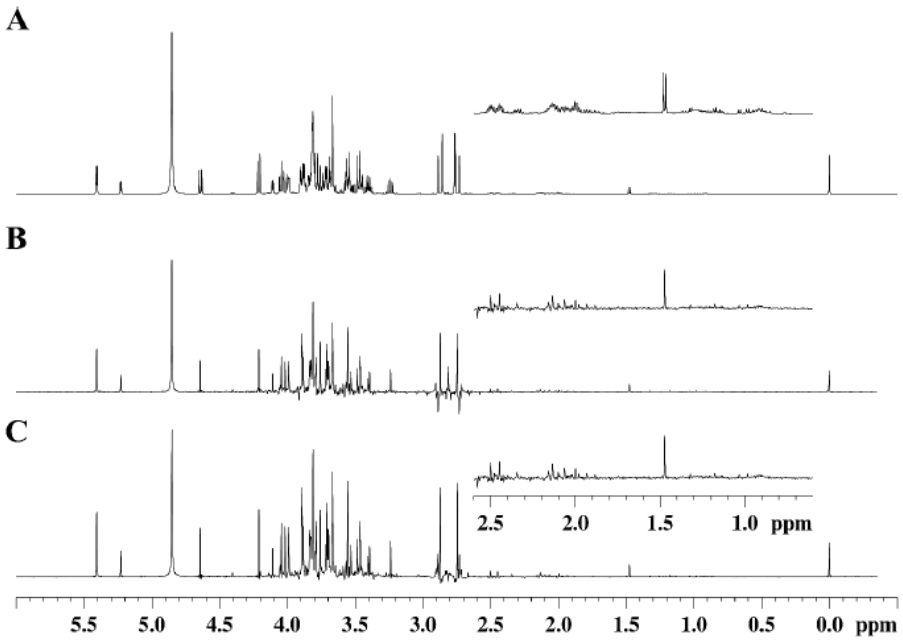


Figure S8. Representative (A) ^{1}H NMR, (B) PSYCHE, and (C) SAPPHIRE-PSYCHE spectra of an aqueous extract of Cape gooseberries (Bambamarca I)

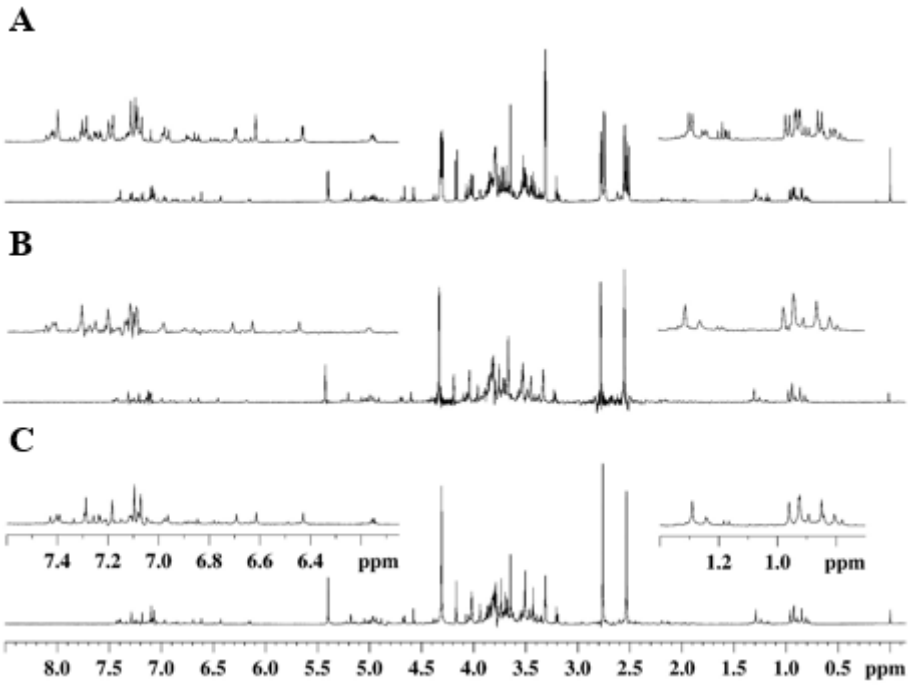


Figure S9. Representative (A) ^{1}H NMR, (B) PSYCHE, and (C) SAPPHIRE-PSYCHE spectra of *Vanilla pompona* extracts.

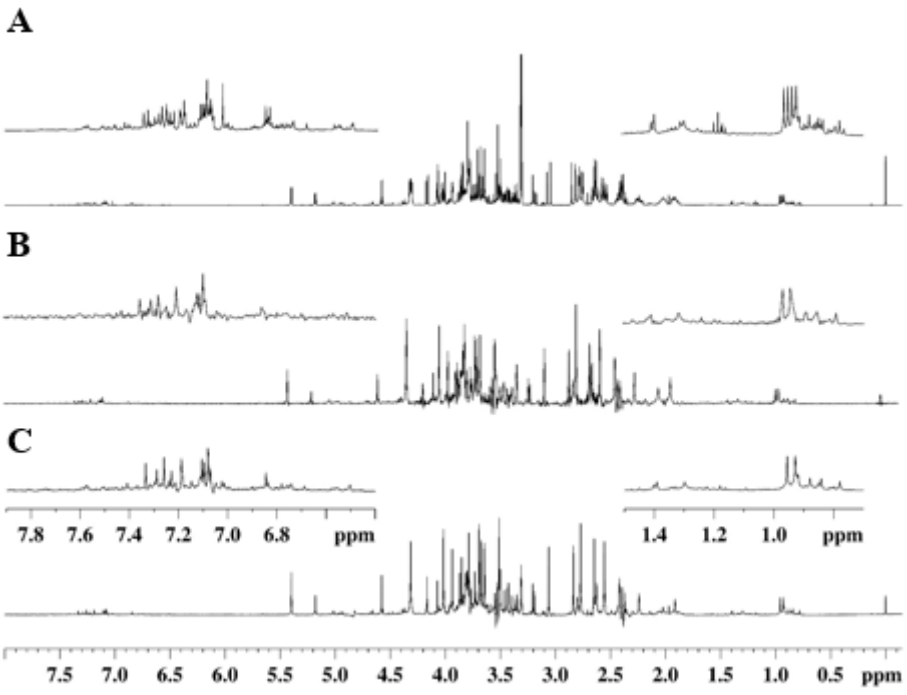


Figure S10. Representative (A) ^{1}H NMR, (B) PSYCHE, and (C) SAPPHIRE-PSYCHE spectra of *Vanilla planifolia* extracts.

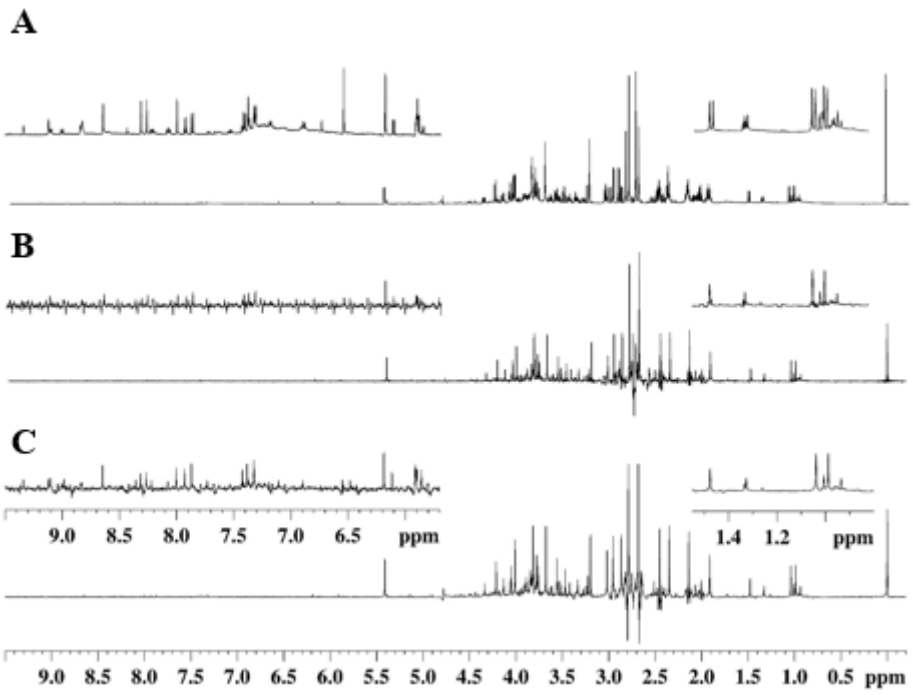


Figure S11. Representative (A) ^{1}H NMR, (B) PSYCHE, and (C) SAPPHIRE-PSYCHE spectra of an aqueous extract of potato.

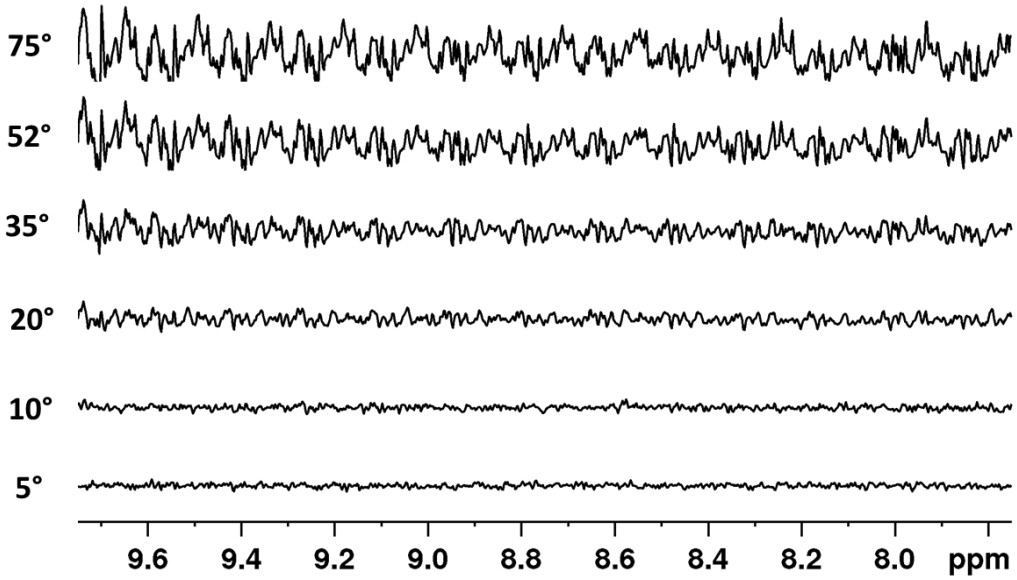


Figure S12. Expanded noise region (from 7.75 ppm to 9.75 ppm) for different PSYCHE spectra of *Vanilla planifolia* extracts using different CHIRP pulse flip angles for the PSYCHE element.