Published in final edited form as:

Anal Chem. 2011 December 15; 83(24): 9352–9360. doi:10.1021/ac201948f.

Measurement of Absolute Concentrations of Individual Compounds in Metabolite Mixtures by Gradient-Selective Time-Zero ^1H - ^{13}C HSQC (gsHSQC₀) with Two Concentration References and Fast Maximum Likelihood Reconstruction Analysis

Kaifeng Hu^{†,*}, James J. Ellinger[‡], Roger A. Chylla[†], and John L. Markley^{†,‡}

National Magnetic Resonance Facility at Madison and Department of Biochemistry, University of Wisconsin-Madison, 433 Babcock Drive, Madison, Wisconsin 53706

Abstract

Time-zero 2D ^{13}C HSQC (HSQC₀) spectroscopy offers advantages over traditional 2D NMR for quantitative analysis of solutions containing a mixture of compounds because the signal intensities are directly proportional to the concentrations of the constituents. The HSQC₀ spectrum is derived from a series of spectra collected with increasing repetition times within the basic HSQC block by extrapolating the repetition time to zero. Here we present an alternative approach to data collection, gradient-selective time-zero ^1H - ^{13}C HSQC₀ in combination with fast maximum likelihood reconstruction (FMLR) data analysis and the use of two concentration references for absolute concentration determination. Gradient-selective data acquisition results in cleaner spectra, and NMR data can be acquired in both constant-time and non-constant time mode. Semi-automatic data analysis is supported by the FMLR approach, which is used to deconvolute the spectra and extract peak volumes. The peak volumes obtained from this analysis are converted to absolute concentrations by reference to the peak volumes of two internal reference compounds of known concentration: DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) at the low concentration limit (which also serves as chemical shift reference) and MES (2-(*N*-morpholino)ethanesulfonic acid) at the high concentration limit. The linear relationship between peak volumes and concentration is better defined with two references than with one, and the measured absolute concentrations of individual compounds in the mixture are more accurate. We compare results from semi-automated gsHSQC₀ with those obtained by the original manual phase-cycled HSQC₀ approach. The new approach is suitable for automatic metabolite profiling by simultaneous quantification of multiple metabolites in a complex mixture.

INTRODUCTION

The primary objective of metabolomics studies is to identify individual chemical components in mixtures and to relate their concentrations to the precise biological state of the system, such as stress, age, and disease.¹ Many methods have been developed to accurately and efficiently identify and profile changes in distinct sets of biomarkers.^{2–4}

*To whom correspondence should be addressed. Address: NMRFAM, Department of Biochemistry, University of Wisconsin Madison, 433 Babcock Drive, Madison, WI 53706-1544, USA. Tel:(608) 263-1722. kaifeng@nmrfam.wisc.edu.

[†]National Magnetic Resonance Facility at Madison.

[‡]Department of Biochemistry.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in the text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Quantitative 1D proton NMR (qHNMR) is used as a routine analytical tool for concentration determination because of its universality, sensitivity, precision, and nondestructive nature.⁵ The integrated intensities of resolved proton resonances in qHNMR spectra are directly proportional to the number of proton spins in the mixture.^{5, 6} However, this approach has shortcomings for signals that are overlapped. 2D ¹H-¹³C HSQC spectra contain a much higher proportion of resolved peaks, but the intensities of cross peaks in conventional 2D NMR experiments are not simply proportional to the concentrations of individual metabolites because of resonance-specific signal attenuation during the coherence transfer periods.⁷ Therefore, conventional 2D NMR experiments usually are not suitable for direct quantitative purpose. Rai et al. presented a method in which signal attenuation factors were calculated for different metabolites and used to evaluate their concentrations from adjusted cross-peak intensities.⁷ An alternative technique is to determine scaling factors from 2D spectra of standards at known concentration collected under the same conditions used for the solution to be analyzed.⁸ We recently reported an approach to acquiring a 2D spectrum (HSQC₀ spectrum) with peak intensities directly proportional to relative concentrations of compounds in mixtures.⁹ The peak intensities in the HSQC₀ spectrum are derived by extrapolating to time zero the peak heights derived from spectra (HSQC_i) collected at the ends of the basic HSQC block repeated once (i=1), twice (i=2), and, optionally, three (i=3) times.⁹ Despite its superior sensitivity, the existence of strong spectral noise, especially *t*₁ noise from metabolites with relatively high concentrations limit the applicability of the original phase-cycled HSQC₀ approach.¹⁰ The *t*₁ noise ridge can obscure cross peaks from metabolites with relatively lower concentrations and thus affect the accuracy of measured peak volumes, and hence reduce the precision and accuracy of concentration estimates. We present here an improved protocol that combines new approaches to data collection and analysis. (1) Phase-cycled spectra are replaced by gradient-selective HSQC_i (gsHSQC_i) spectra, which results in much cleaner spectra (Figure S1). (2) Semi-automated data analysis is carried out by the fast maximum likelihood reconstruction (FMLR) approach¹¹ as implemented in the Newton software package.¹² (3) Two internal reference compounds that span the concentration range of the compounds being analyzed are added to the mixture to be analyzed, which improves the precision and accuracy of compound quantification. We anticipate that this new protocol for 2D ¹H-¹³C spectroscopy will prove valuable for automatic metabolite profiling by simultaneous quantification of multiple metabolites in a complex mixture.

METHODS

Pulse Sequence Design

In implementing gsHSQC₀ spectroscopy to suppress *t*₁ artifacts, different selective gradient pairs must be used in each HSQC_i block to avoid gradient refocusing among the HSQC_i units. The selective gradients must follow: $\Delta G_{12} = \Delta G_{34} = \Delta G_{56} = 0$; $\Sigma G_{12} \neq \Sigma G_{34} \neq \Sigma G_{56}$; and $\Sigma G_{12} \pm \Sigma G_{34} \neq \Sigma G_{56}$ (See the theoretical description of gsHSQC_i in the Supporting Information). Differences in the signal attenuation factors among the HSQC blocks due to the application of different selective gradient pairs are small and can be neglected. We have found that equivalent peak volumes are obtained even though HSQC_i data are acquired with the different sets of selective gradients as illustrated in Figure 1. The loss of sensitivity in gsHSQC_i due to gradient selection is partially compensated by running gsHSQC_i as the non-constant time version (Figure 1C).

In quantitative, gradient-selective 2D ¹H-¹³C HSQC (Figure 1), immediately after the first 90° excitation pulse (with phase *x*), the density operator I_{-y} is generated, which can be decomposed as a linear combination of $\frac{1}{2} I_{-}$ and $\frac{1}{2} I_{+}$. The density operator right before detection (point *f*) has the same format as the density operator generated immediately after the first 90° excitation pulse, which can be decomposed as linear combination of $\frac{1}{4} I_{-}$ and $\frac{1}{4} I_{+}$.

I_+ , of which only the $\frac{1}{4} I_-$ term is detected by the spectrometer. Compared to the density operator immediately after the 90° excitation pulse ($\frac{1}{2} I_-$ and $\frac{1}{2} I_+$), the uniform scaling factor of $\frac{1}{2}$ (from $\frac{1}{2} I_-$ to $\frac{1}{4} I_-$ and from $\frac{1}{2} I_+$ to $\frac{1}{4} I_+$) is due to the gradient selection during the coherence pathway. A detailed description of the relevant coherence pathways is given in the Supporting Information. Figure 1B shows the scheme of the relevant coherence pathway for echo (indicated by the solid arrow lines in upper scheme) and antiecho (indicated by the dashed arrow lines in the lower scheme), respectively. In each situation, the coherence pathways in red and green are both selected, but only the $\frac{1}{4} I_-$ density operator term (green coherence pathway) will be detected by the spectrometer. In both approaches, an amplitude attenuation factor $f_{A,n}$ specific for each functional group arises during the coherence pathway due to relaxation, imperfect pulses, mismatch of the INEPT transfer delay with J -coupling, etc.⁹ Thus, the amplitudes of the detected signals for phase-cycled approach are given by Eq 1, whereas those for the gradient-selective approach are given by Eq 2:

$$A_{1,n}(I_-) = A_{0,n}(I_-) \cdot f_{A,n} \quad (1)$$

$$A_{1,n}(I_-) = 2A_{0,n}(I_-) \cdot (f_{A,n} \cdot \frac{1}{2}) \quad (2)$$

in which $A_{1,n}(I_-)$ is the integrated peak intensity (signal amplitude) of peak n ; $A_{0,n}$ is the virtual integrated signal intensity (signal amplitude) of peak n immediately after the first 90° excitation pulse; and $f_{A,n}$ is the amplitude attenuation factor specific for peak n . In the gradient-selective approach (Eq 2) the scaling factor of $\frac{1}{2}$ arises from the gradient selection, and the factor of 2 in front of $A_{0,n}$ accounts for signal enhancement resulting from the linear recombination of the time domain NMR data recorded in the echo-antiecho mode. The $f_{A,n}$ factor accounts for signal losses during the coherence transfer periods from point a (Figure 1A, immediately after the first ^1H excitation pulse) to point f (Figure 1A, immediately before acquisition). This attenuation factor $f_{A,n}$ is specific to a particular cross peak (peak n) because of different chemical environments, dynamics, relaxation properties, and J -couplings. Therefore, the peak intensity $A_{1,n}(I_-)$ in 2D ^{13}C gsHSQC is not directly proportional to the number of spins giving rise to the signals. However, as in phase-cycled HSQC,⁹ the scaling factor $f_{A,n}$ can be determined simply by inserting the pulse sequence components included in the brackets (Figure 1D) in front of the HSQC₁ unit. Three 2D HSQC_{*i*} ($i = 1, 2, 3$) spectra are acquired, in which the subscript i indicates the number of times the basic building block is repeated. Note that in HSQC₂ and HSQC₃, additional phase cycling is applied on ϕ_3 and ϕ_4 and on the receiver phase ϕ_{rec} . Figure 1E shows the relevant coherence pathway for the full pulse sequence of gradient-selective 2D HSQC_{*i*} ($i=1, 2, 3$). Echo (indicated by the solid arrow lines in upper scheme) and antiecho (indicated by the dashed arrow lines in the lower scheme) quadrature detection is achieved in the HSQC₁ unit. For both situations, both coherence pathways in red and green are selected, but only the $1/16 I_-$ density operator term (green coherence pathway) is detected. In HSQC₂ and HSQC₃ units (Figure 1E), spin “echo” gradient selection (coherence order in black and corresponding solid arrow lines in both upper and lower scheme) is achieved through selective gradient pairs, g_3/g_4 and g_5/g_6 , respectively. The density operators detected in 2D gsHSQC_{*i*} ($i=1, 2, 3$) all have the same format, which can be decomposed as linear combination of $\frac{1}{4} I_-$ and $\frac{1}{4} I_+$; $1/8 I_-$ and $1/8 I_+$; $1/16 I_-$ and $1/16 I_+$, respectively (Figures 1D and 1E). The virtual peak intensity of peak n in gsHSQC₀, $A_{0,n}$, can be obtained through linear regression extrapolation of:

$$\ln(A_{i,n}) = \ln(2A_{0,n}) + i \times \ln(f_{A,n} \cdot \frac{1}{2}) \quad (3)$$

whereas the virtual peak intensity of peak n in phase-cycled HSQC₀ is obtained through linear regression extrapolation of: ⁹

$$\ln(A_{i,n}) = \ln(A_{0,n}) + i \times \ln(f_{A,n}) \quad (4)$$

Sample Preparation

Model mixture of metabolites: L-alanine, L-methionine, 3-hydroxybutyrate and D-glucose were purchased from Sigma-Aldrich. 2-(N-morpholino)ethanesulfonic acid (MES) was from Research Organics. DSS and 99.8% D₂O were from Cambridge Isotope Laboratories. All dry reagents were stored under desiccation to ensure that they were free of moisture absorbed from the air. Alanine, methionine, 3-hydroxybutyrate, D-glucose and MES were weighed directly into a grade A 50 mL Kimax volumetric flask to give final concentrations of 12.35, 18.14, 26.78, 22.32 and 49.93 mM, respectively. A 35.57 mM DSS stock solution was prepared by weighting. Aliquots of the stock solution were added to volumetric flasks to give final concentration of 3.557 mM (1:10 dilution). Each reagent was prepared from weighing to at least 3 significant figures on an analytical balance of ± 0.1 mg precision. Approximately 40 mL of D₂O was slowly added to the sample while gently swirling; the observed pH was adjusted to 7.400 using deuterated base; and the final volume was adjusted to 50 mL. The solution rested at least 6 h to ensure anomeric equilibration of D-glucose. Prior to NMR data acquisition, 12 μ L of 100 mM Fe(III)EDTA was added to 500 μ L of the mixture as a relaxation agent; with this sample, the longest proton T_1 was measured to be shorter than 0.433 s. Bovine liver extract: Bovine liver was purchased from a local organic grocer, frozen, and lyophilized until dry. After lyophilization, a morsel (420 mg) was macerated gently with a metal spatula in a mortar placed on dry ice. The tissue was transferred to a 50 mL conical tube on dry ice and homogenized with a rounded glass rod. Then, 16 mL hot (95° C), distilled, deionized water was added, and the sample was briefly agitated by vortex mixing. The sample was then placed in a 95° C water bath for 7.5 min. Samples were removed from the hot water bath, briefly agitated by vortex mixing, chilled on ice, and centrifuged at 8000 RPM for 5 min (4° C) to remove cellular debris. The resulting supernatant was filtered through a 3 kDa cut-off ultrafiltration membrane. The filtered liver extract was frozen, lyophilized, and resuspended in 1 mL 99.9% D₂O. The final sample was prepared by mixing 786 μ L of extract with 200 μ L MES (249.65 mM stock solution) and 14 μ L DSS (35.57 mM stock solution). The final concentrations of MES and DSS were 49.93 mM and 0.498 mM, respectively, which served as concentration references for quantification. Finally, the sample was titrated to a pH reading of 7.350 by adding a small amount of 1 M NaOD. 600 μ L of the above sample was transferred to an NMR tube and used directly to collect HSQC_i data. A relaxation enhancing agent, 14 μ L of 100 mM Fe(III)EDTA, was added to a second 600 μ L sample of the liver extract to shorten the inter-scan delay used in HSQC_i data collection.

NMR Experiments

For the model mixture of metabolites, NMR data were collected on a 700 MHz Bruker Avance III spectrometer equipped with a QCI probe, with radio-frequency pulses applied on ¹H and ¹³C at 4.8 ppm and 48 ppm, respectively. GARP ¹³C decoupling used a field strength of $\gamma B_2 = 2.5$ kHz. 2048 $\times$ 75 complex data points with spectral width of 16 ppm and 80 ppm, respectively, were collected along the ¹H and ¹³C dimensions, with 16 scans per

FID and an inter-scan delay of 2.5 s (longer than 5 times of the longest proton T_1 value), resulting in a total acquisition time of 2 h for each HSQC_i.

The constant-time delay T was set to 10.4 ms. Constant-time gsHSQC_i data were collected with two different sets of selective gradients. (Set 1) g_1 : 42.4 G/cm (80%); g_2 : 10.653 G/cm (20.1%); g_3 : 37.1 G/cm (70%); g_4 : 9.3214 G/cm (17.5875%); g_5 : 31.8 G/cm (60%); and g_6 : 7.98975 G/cm (15.075%) (see Supporting Information, Tables S1 and S5). (Set 2) g_1 : 42.4 G/cm (80%); g_2 : 10.653 G/cm (20.1%); g_3 : 31.8 G/cm (60%); g_4 : 7.98975 G/cm (15.075%); g_5 : 21.2 G/cm (40%); and g_6 : 5.3265 G/cm (10.05%) (see data Table S2 and S6). For comparison, constant-time phase-cycled HSQC_i (see Supporting Information, Tables S3 and S7) and non-constant time gsHSQC_i (see Supporting Information, Tables S4 and S8) data were also collected.

For the bovine liver extract with and without added relaxation enhancing agent, constant-time gsHSQC_i data were collected with gradient set 1 (above) on a 600 MHz Bruker Avance III spectrometer equipped with a QCI probe by setting an inter-scan delay of 2.0 s and 11.0 s, respectively. 2048×80 complex data points with spectral width of 16 ppm and 80 ppm, respectively, were collected along the ^1H and ^{13}C dimensions. For bovine liver extract without relaxation enhancing agent, 16 scans per FID are acquired resulting in a total acquisition time of 8 h for each HSQC_i; for the sample with Fe(III)EDTA, 32 scans per FID were acquired resulting in a total acquisition time of 3 h for each HSQC_i.

Peak Amplitude and Peak Intensity Measurements

Spectral deconvolution of data sets was performed using the method of fast maximum likelihood reconstruction (FMLR)¹¹ as implemented in the Newton software package (Newton-FMLR).¹² Newton-FMLR iteratively constructs a time-domain model with the fewest number of signals whose Fourier transformation matches the peaks visible in the spectrum. The HSQC₁ data set was used as a reference data set for deconvolution. The signal model obtained from this data set was propagated to the HSQC₂ and HSQC₃ data sets, where the amplitudes of the signals were determined by linear least squares analysis. In cases where more than one peak (signal multiplicity) is associated with a given atom, a region of interest was constructed, and the sum of the amplitudes in the region was used to perform the extrapolation of the time zero amplitude.¹² For the model mixture of metabolites, the sum of the amplitudes of the reconstructed signals is analogous to integration of the spectral region (Supporting Information, Tables S5–S8). For comparison, the peak intensity was also obtained through manual integration (Supporting Information, Tables S1–S4) as described previously.⁹ For bovine liver extract samples, 47 regions of interest (ROI) (Figure 4) were chosen for quantification of 23 metabolites in the mixture. The signal amplitudes were obtained by fast maximum likelihood reconstruction (FMLR) analysis (Tables S9–S10 for bovine liver extract samples with and without relaxation enhancing agent).

Absolute Concentrations of Individual Metabolites

The measured peak volumes were related to the absolute concentrations by using two compounds with known gravimetric concentrations (C_g) as concentration references: DSS (which also served as the chemical shift reference) at the low concentration limit and MES at the high concentration limit. With two concentration references, the linear relation between the concentration and the obtained peak volume is defined by: $C_g = kA_0 + b$, in which k and b are calculated from the averaged normalized peak intensity or peak amplitude (A_0) of DSS and MES and their corresponding gravimetric concentrations (Supporting Information, Tables S1A–S8A, for the model mixture of metabolites; Table S11, for the bovine liver extract samples, relevant information is highlighted in blue). The non-zero

constant b may indicate a non-zero baseline offset (gray background). On the other hand, b can be set to 0 to force the concentration to be linearly proportional to the obtained peak intensities or peak amplitudes (A_0). Once the values of k and b are obtained, the measured concentration of each metabolite is then calculated as $C_m = kA_0 + b$ from the measured normalized peak intensity or peak amplitude A_0 of each individual cross peak. The accuracy and precision of the measured absolute concentration of each individual chemical in the metabolite mixture was analyzed (Supporting Information, Tables S1A–S8A).

RESULTS

Figure 2 shows the 2D gradient-selective constant-time HSQC₁ spectrum of the model metabolite mixture containing DSS, alanine (A), methionine (M), 3-hydroxybutyrate sodium (HB), D-glucose (G) and MES in D₂O, in which there is thermodynamic equilibrium between the α -D-glucose (α G) and β -D-glucose (β G) at 36%:64%. The peaks are labeled with the molecule identifier and the position of the functional group. The manually integrated peak intensities for these compounds are shown in the Supporting Information (Tables S1–S4), and the peak amplitudes extracted by Newton-FMLR from the HSQC₁ are shown in the Supporting Information (Tables S5–S8). The measured concentration of each metabolite calculated using the manually integrated peak intensity and peak amplitude A_0 from Newton-FMLR are shown in the Supporting Information (Tables S1A–S4A and Tables S5A–S8A, respectively).

Equations 3 and 4 predict a difference of $\ln(2)$ between the intercepts in the linear regression of gsHSQC _{i} and phase-cycled HSQC _{i} data. Using both manual integrated peak volumes (Figure 3A) and peak amplitudes from semi-automated Newton-FMLR analysis (Figure 3B), this difference can be seen in the analyzed results for gsHSQC _{i} (Figure 3: (1) and (2) solid lines with selective gradient sets 1 and 2, respectively.) and phase-cycled HSQC _{i} (Figure 3: (3) States-TPPI_{CT}, dashed lines). The uniform scaling factor of $\frac{1}{2}$ due to the gradient selection can lead to the quick signal attenuation in gsHSQC _{i} as i increases. This can be partially alleviated by collecting gsHSQC _{i} data in non-constant-time mode (Figure 1C), which yields a larger cross peak specific amplitude attenuation factor $f_{A,n}$ than in constant-time mode (Figure 3: (4), dotted lines).

Compared to the constant-time gsHSQC _{i} (Figure 3: (1) and (2) solid lines), linear regression with the non-constant-time gsHSQC _{i} (Eq. 3) gives a slightly shallower slope (Figure 3: (4) dotted lines) due to the slower signal attenuation as i increases. Note that although the cross peak specific amplitude attenuation factor $f_{A,n}$ is different for constant-time and non-constant-time gsHSQC _{i} , linear regression with both constant-time and non-constant-time gsHSQC _{i} results in the same $A_{0,n}$ value, that is, the same virtual gsHSQC₀ (Figure 3: (1), (2), and (4) have the same intercepts).

The accuracy and precision of the obtained absolute concentration for individual chemicals in the model metabolite mixture are analyzed in Supporting Information (Tables S1A–S8A). Table 1 summarizes the average accuracy and precision of the measured concentration as well as the obtained percentage of α -D-glucose by different approaches: constant-time gsHSQC₀ with two different sets of selective gradients, non-constant-time gsHSQC₀, and phase-cycled HSQC₀ as analyzed by manual integration (direct summation) of peak volumes and by semi-automated Newton-FMLR. The accuracy and precision of the results obtained by semi-automated Newton-FMLR are in agreement with concentrations obtained by manual integration (Table 1).

To further explore the performance of the approach, we applied the gsHSQC₀ together with semi-automated fast maximum likelihood reconstruction (FMLR) data analysis to quantify

individual metabolites in a bovine liver extract. Figure 4 shows the 2D gradient-selective constant-time HSQC₁ spectrum of the bovine liver extract containing the relaxation enhancing agent, Fe(III)EDTA.

47 ROIs (Figure 4) were chosen for quantification of 23 metabolites in the mixture. Signal amplitudes obtained from Newton-FMLR analysis of constant-time gsHSQC₁ of bovine liver extract with and without Fe(III)EDTA are shown in Table S9 and S10, respectively. Using DSS and MES at respective gravimetric concentrations (C_g) of 0.498 mM and 49.930 mM as concentration references and the linear relation between the concentration and the normalized peak volume A_0 , $C_g = kA_0 + b$, yielded values for k and b (Table S11). Measured concentrations of metabolites in the bovine liver extract with or without Fe(III)EDTA were then calculated from individual normalized peak volumes A_0 using the known k and b values, $C_m = kA_0 + b$ (Table 2). The results obtained from the bovine liver extract sample with the relaxation enhancing agent are in agreement with those without the relaxation enhancing agent, but the latter required much longer acquisition times as described in NMR experiments. Moreover, without the relaxation agent, peaks in the spectra of the bovine liver extract sample from pantothenate methyl groups C_a and C_b and methionine C_γ were too weak to be detected and used for quantitative purpose. The relatively poor precision for pantothenate and taurine (Table 2, grey background) is probably due, respectively, to overlaps of signals from the pantothenate C_a and leucine $C_{\delta 2}$ methyl group and to the effect of the signal from the C_1 group of MES (present at high concentration) on the relatively weak signal from the C_1 of taurine.

DISCUSSION

Gradient-selective HSQC_i (gsHSQC_i) led to cleaner spectra and the partial removal of t_1 noise ridges (Supporting Information, Figure S1). The strength of the t_1 noise usually is proportional to the intensity of related peaks. Therefore, the measured concentrations for compounds present at high concentration in the mixture usually are less precise than those from compounds present at moderate concentrations. By using a content-defined model mixture of metabolites to evaluate the performance of the approach, (Table 1), comparable precision and accuracy were obtained by manual integration and by semi-automated Newton-FMLR analysis.¹² The gsHSQC₀ approaches (both constant-time and non-constant-time) yielded better precision, but no better accuracy than phase-cycled HSQC₀. Data analyzed by Newton-FMLR showed better accuracy with two-parameter (k and b) fitting than with one-parameter (k) fitting, but comparable precision was obtained for both cases. With manual integration, the two fitting approaches yielded equivalent results.

As explained in the Supporting Information, to avoid quadrature imaginary artifacts in gsHSQC₂ and gsHSQC₃ spectra, which arise from gradient refocusing among HSQC units, different selective gradient pairs must be used in each HSQC block. Using different selective gradient pairs in each HSQC_i block could cause the signal attenuation factors $f_{A,n}$ to be different in each HSQC_i block; however, our experimental results revealed that this effect is negligible (Supporting Information, Tables S1 and S2). Similar gsHSQC₀ peak volumes and calculated concentrations were obtained from gsHSQC_i data acquired with different sets of selective gradients (Figure 3 (1) and (2)).

Concentrations were determined for all components in the mixture by manual integration (direct summation) of peak volumes and by Newton-FMLR analysis to extract the peak amplitudes in HSQC_i spectra.¹² These concentrations were found to be comparable (Supporting Information, Tables S1A–S8A), which shows that the semi-automated approach should be applicable to quick metabolite profiling.

To explore the performance of the approach on a real biological metabolite sample, we applied the gsHSQC₀ with semi-automated FMLR data analysis to quantify individual metabolites in a bovine liver extract. Concentrations of 23 metabolites were determined. Overall, concentrations determined from the samples with and without relaxation enhancing agent were in agreement, but the latter sample required a much longer acquisition time. With the acquisition time for each gsHSQC_i of 3 h for the bovine liver extract sample with relaxation enhancing agent, the lower limit of measurable concentration was on the order of 0.1 mM. For example, the concentration of betaine was 0.2–0.3 mM, and the concentration of methionine was 0.4–0.5 mM. Glucose was the compound with the highest concentration: 120 mM for the sum of the alpha and beta anomers.

We recently applied the non-constant-time gsHSQC₀ approach to selectively quantify a natural product of thiocoraline present at a level of 1% w/w (equivalent concentration of the pure compound ~0.1 mM) in a complex extract mixture from a *Verrucosipora* sp. isolated from the sponge *Chondrilla caribensis* f. *caribensis*.¹³

The use of gradient selection can greatly suppress spectral noise, especially t_1 noise from compounds present at relatively high concentrations in the mixture. Compared to the phase-cycled HSQC₀ approach,⁹ the signal amplitudes in gsHSQC_i attenuate faster by the factor of $\frac{1}{2}$ due to the gradient selection (Eq 3 and 4). This is reflected in the generally more negative slopes of the linear regression plots of $\ln A_{i,n}$ vs. i (Figure 3). Thus, signals from very dilute compounds may decay and become very weak in the HSQC₃ spectrum. In such cases, it is better to acquire only two spectra (HSQC₁ and HSQC₂) and to determine the $A_{0,n}$ values from:

$$2A_{0,n} = A_{1,n}^2 / A_{2,n} \quad (6)$$

CONCLUSIONS

We recently proposed ¹H-¹³C zero-time HSQC (HSQC₀)⁹ spectroscopy as a way to extend quantitative NMR from 1D proton (1D qHNMR) to two dimensions in order to improve peak resolution and quantification. We next extended the approach to make use of two internal concentration references and semi-automated analysis by Newton-FMLR.¹² Our original approach utilized a phase-cycled constant-time HSQC pulse sequence. Here we evaluate three additional pulse sequences and compare them with the original. The new approaches utilize gradient selection for quadrature detection in place of phase-cycled quadrature detection: (1) gradient-selective constant-time with gradient set 1, (2) gradient-selective constant-time with gradient set 2, and (3) gradient-selective non-constant-time. Gradient selection yields HSQC_i spectra with fewer spectral artifacts and can improve the precision of peak amplitude measurements for more intense peaks. Apart from this, data from all four approaches are amenable to analysis either by manual integration or by semi-automated Newton-FMLR analysis,¹² and all four yield concentration measurements of comparable accuracy. The performance of the approach on a real biological metabolite sample was evaluated by analyzing a bovine liver extract. We collected gsHSQC₀ data and analyzed the results with semi-automated FMLR to determine the concentrations of 23 metabolites present in the extract. In this case, the lower limit of measurable concentration was on the order of 0.1 mM and the highest concentration measured was about 120 mM. We anticipate that this new protocol for 2D ¹H-¹³C spectroscopy will prove valuable for automatic metabolite profiling by simultaneous quantification of multiple metabolites in a complex mixture.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the NIH National Center for Research Resources (grants P41 RR02301 and 3P41 RR02301-25S1) and by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-07ER64494).

References

1. Wishart DS. TrAC-Trends Anal Chem. 2008; 27:228–237.
2. Piotto M, Moussallieh FM, Dillmann B, Imperiale A, Neuville A, Brigand C, Bellocq JP, Elbayed K, Namer IJ. Metabolomics. 2009; 5:292–301.
3. Zhang F, Dossey AT, Zachariah C, Edison AS, Bruschweiler R. Anal Chem. 2007; 79:7748–7752. [PubMed: 17822309]
4. Zhang FL, Bruschweiler R. Angew Chem Int Edit. 2007; 46:2639–2642.
5. Pauli GF, Jaki BU, Lankin DC. J Nat Prod. 2005; 68:133–149. [PubMed: 15679337]
6. Akoka S, Barantin L, Trierweiler M. Anal Chem. 1999; 71:2554–2557. [PubMed: 21662801]
7. Rai RK, Tripathi P, Sinha N. Anal Chem. 2009; 81:10232–10238. [PubMed: 19919088]
8. Lewis IA, Schommer SC, Hodis B, Robb KA, Tonelli M, Westler WM, Suissman MR, Markley JL. Anal Chem. 2007; 79:9385–9390. [PubMed: 17985927]
9. Hu KF, Westler WM, Markley JL. J Am Chem Soc. 2011; 133:1662–1665.
10. Poulding S, Charlton AJ, Donarski J, Wilson JC. J Magn Reson. 2007; 189:190–199. [PubMed: 17920317]
11. Chylla RA, Volkman BF, Markley JL. J Biomol NMR. 1998; 12:277–297. [PubMed: 9751999]
12. Chylla RA, Hu K, Ellinger JJ, Markley JL. Anal Chem. 2011; 83:4871–4880. [PubMed: 21526800]
13. Hu KF, Wyche TP, Bugni TS, Markley JL. J Nat Prod. 2011

Figure 1.

(A) Basic building block of pulse sequence of the gradient selective constant time 2D HSQC for quantification. Narrow and wide black bars indicate 90° and 180° pulses, respectively. The delays are $\tau = 3.45$ ms; $T = 10.4$ ms. The phase cycling is as follows: $\phi_1 = x, -x$; $\phi_2 = 2[x], 2[-x]$; $\phi_{\text{rec}} = x, -x, -x, x$. All other radio-frequency pulses are applied with phase x except as indicated. Quadrature detection in the ^{13}C (t_1) dimension is achieved using echo-antiecho through flipping the polarity of gradient g_1 (filled and dashed sine bell). The 180° pulse marked with asterisk is inserted to refocus chemical shift of proton during the gradient delays for g_{13} and g_{14} to achieve better t_1 noise suppression. All pulsed field gradients are applied along the z -axis with a duration of 1 ms followed by gradient recovery period of 200 μs . The strength of the gradients (percentage of highest gradient current, 10 A) are: g_{13} : 37.1 G/cm (70%); g_{14} : 45.05 G/cm (85%); g_{15} : 7.95 G/cm (15%), and g_{16} : 47.7 G/cm (90%), and the selective gradient pairs are g_1 : 42.4 G/cm (80%) and g_2 : 10.653 G/cm (20.1%). (B) Diagram showing the relevant coherence pathways for echo (indicated by the solid arrow lines in upper scheme) and antiecho (indicated by the dashed arrow lines in the lower scheme). In both situations, both the coherence pathways shown in red and green are selected, but only the I_- density operator terms (green coherence pathway) are detected by the spectrometer. (C) Basic building block of the gradient-selective non-constant time 2D HSQC pulse sequence used for quantification. The short delay $\delta_1 = 1.21$ ms is just to

accommodate the coherence selective pulsed field gradients g_1 , including the gradient recovery period of 200 μs . **(D)** Full gradient-selective 2D HSQC_{*i*} (*i* = 1, 2, 3) pulse sequence. HSQC units HSQC₂ and HSQC₃ with corresponding selective gradient pairs (g_3/g_4 and g_5/g_6 , respectively) are inserted in front of HSQC₁, in which chemical shift of ^{13}C (t_1) is allowed to evolve. Within HSQC₂ and HSQC₃ units, the constant time period T is replaced by the invariant constant time T without encoding chemical shift evolution. The additional phase cyclings for HSQC₂ are $\varphi_3 = 4[x]$, $4[-x]$, $\varphi_{\text{rec}} = x, -x, -x, x, -x, x, x, -x$ and for HSQC₃ are $\varphi_4 = 8[x]$, $8[-x]$, $\varphi_{\text{rec}} = x, -x, -x, x, -x, x, x, -x, -x, x, x, -x, x, -x, -x, x$. To avoid quadrature imaginary peaks due to gradient echoes between HSQC blocks, different selective gradient pairs must be used in each HSQC block. Differences in signal attenuation due to different selective gradient pairs applied among the HSQC blocks can be ignored (Figure 3 (1) and (2), overlapped). It is demonstrated by using two different sets of selective gradients as: Set 1) g_1 : 80%; g_2 : 20.1%; g_3 : 70%; g_4 : 17.59%, g_5 : 60%; g_6 : 15.075% of the maximum of 53 G/cm. (Data Table S1 and S5). Set 2) g_1 : 80%; g_2 : 20.1%; g_3 : 60%; g_4 : 15.075%, g_5 : 40%; g_6 : 10.05% of the maximum of 53 G/cm. (Data Table S2 and S6). **(E)** Diagram showing the relevant coherence pathways of **(D)**, with the echo coherence pathway indicated by the solid arrow lines (upper scheme) and the antiecho indicated by the dashed arrow lines (lower scheme). Quadrature detection is achieved in the HSQC₁ unit. In HSQC₂ and HSQC₃, spin “echo” gradient selection (coherence order in black and corresponding solid arrow lines in both upper and lower schemes) is achieved through selective gradient pairs g_3/g_4 and g_5/g_6 , respectively.

Figure 2.

700 MHz 2D gradient selective constant time HSQC₁ spectrum of the model metabolite mixture of 12.35 mM alanine (A), 18.14 mM methionine (M), 26.78 mM 3-hydroxybutyrate sodium (HB), and 22.32 mM D-glucose (G) in D₂O containing 3.557 mM DSS and 49.93 mM MES as concentration references. The glucose exists as a thermodynamic equilibrium between 36% α-D-glucose (αG) and 64% β-D-glucose (βG). The molecular structures and corresponding naming of these compounds are shown. The peaks are labeled with the molecular abbreviated name and the position of the functional group.

Figure 3.

Extrapolation of methionine methyl ($M:CH_3$) 1H - ^{13}C cross peak intensities from $HSQC_i$ data acquired at 700 MHz (1H) by (1) constant-time gradient-selective $HSQC_i$ (gs $HSQC_i$) with selective gradient set 1 (EA_CT_grad_set1), (2) gs $HSQC_i$ with selective gradient set 2 (EA_CT_grad_set2), (3) constant-time phase-cycled $HSQC_i$ (States-TPPI_CT), and (4) non-constant-time gradient-selective $HSQC_i$ with selective gradient set 1 (EA_non-CT). Peak volumes were determined by (A) manual integration and (B) semi-automated Newton-FMLR analysis. Note line (1) is covered under that of (2).

Figure 4. 600 MHz 2D constant-time gradient-selective HSQC₁ spectrum of bovine liver extract sample with relaxation enhancing agent, Fe(III)EDTA in D₂O containing 0.498 mM DSS and 49.93 mM MES as concentration references. 47 labeled peaks were chosen for quantification of 23 metabolites in the mixture.

Table 1

Accuracy and precision of the measured concentrations and obtained percentage of α D-glucose from peak volumes obtained by manual integration and semi-automated Newton-FMLR analysis from four extrapolated time-zero HSQC approaches.

Manual integration		Constant-time gsHSQC ₀ , set 1		Constant-time gsHSQC ₀ , set 2		Non-constant-time gsHSQC ₀		Phase-cycled HSQC ₀	
% α -D-Glucose		38.35%	38.47%	36.23%	36.75%	38.01%	38.20%	37.62%	37.88%
Average accuracy		5.50%	4.81%	1.97%	4.33%	5.40%	4.69%	2.81%	3.23%
Average precision		3.70%	3.68%	5.52%	5.40%	3.61%	3.56%	7.52%	7.39%
Newton-FMLR		Constant-time gsHSQC ₀ , set 1		Constant-time gsHSQC ₀ , set 2		Non-constant-time gsHSQC ₀		Phase-cycled HSQC ₀	
% α -D-Glucose		38.66%	38.27%	37.53%	38.27%	36.04%	37.00%	37.84%	38.32%
Average accuracy		4.89%	8.95%	3.70%	9.40%	4.50%	10.82%	3.20%	5.25%
Average precision		4.86%	4.64%	6.11%	5.72%	6.62%	6.03%	8.51%	8.11%

Values were calculated for k and b , by using the linear relation between the gravimetric concentration C_g and the normalized peak volume A_0 ($C_g = kA_0 + b$) and the known gravimetric concentrations of DSS and MES. The parameter b was set as 0 (white background) or was allowed to float to account for non-zero baseline offset (grey background).

Measured concentrations were then calculated from individual A_0 values and the known k and b values by the equation, $C_m = kA_0 + b$, in which b is 0 (white background) or non-zero intercept (grey background).

The average concentration was calculated as the average of C_m values for cross peaks belonging to the same individual chemical. Stdev was calculated as the standard deviation of C_m values for cross peaks belonging to the same individual chemical.

The percentage of α -D-glucose was calculated as: Ave. conc. of α G / (Ave. conc. of α G + Ave. conc. of β G).

$$\text{Accuracy was calculated as: } \sqrt{\frac{\sum_i^n (\text{Precision}_i)^2}{n}} \quad ; \quad \text{Ave. Precision was calculated as: } \sqrt{\frac{\sum_i^n (\text{Accuracy}_i)^2}{n}} \quad ; \quad \text{Ave. Accuracy was calculated as: } \frac{\sum_i^n (\text{Ave. conc.} \cdot C_m - C_g) / C_g}{\sum_i^n (\text{Ave. conc.} \cdot C_m - C_g) / C_g}$$

Table 2

Measured concentrations of metabolites in the bovine liver extract with or without Fe(III)EDTA using constant-time gsHSCQ_i

Compound: 2D NMR cross peak used in quantitation	Liver_extract with Fe(III)EDTA				Liver_extract without Fe(III)EDTA			
	Conc. Calcd. <i>k_b</i> (mM)	Ave./Precision	Conc. Calcd. <i>k_b</i> (mM)	Ave./Precision	Conc. Calcd. <i>k_b</i> (mM)	Ave./Precision	Conc. Calcd. <i>k_b</i> (mM)	Ave./Precision
MES:C4-H4	47.341	49.930	47.344	49.929	48.923	49.930	48.923	49.930
MES:C2-H2	50.920	3.70%	50.919	3.69%	50.630	2.14%	50.631	2.14%
MES:C3-H4	49.932		49.931		49.118		49.118	
MES:C1-H2 ^o	51.527		51.524		51.048		51.049	
DSS:Met-H9	0.498	0.498	0.556	0.556	0.498	0.498	0.476	0.476
3-hydroxybutyrate:CH3	0.248	0.248	0.307	0.307	0.637	0.637	0.615	0.615
Acetate:CH3	0.927	0.927	0.985	0.985	1.170	1.170	1.149	1.149
Ala:Cα-H1	5.514	5.244	5.566	5.296	6.203	5.917	6.184	5.898
Ala:Cβ-H3	4.974	7.28%	5.027	7.20%	5.631	6.83%	5.612	6.86%
Asp:Cα-H1	0.849	0.849	0.907	0.907	1.316	1.316	1.294	1.294
Betaine:Met-H9	0.232	0.232	0.290	0.290	0.136	0.136	0.114	0.114
Camitine:Met-H9	7.930	7.930	7.979	7.979	10.110	10.110	10.093	10.093
Choline:C1-H2	3.789	3.852	3.843	3.906	4.749	4.661	4.729	4.641
Choline:C2-H2	3.557	8.58%	3.612	8.45%	4.596	1.70%	4.576	1.71%
Choline:Met-H9	4.209		4.263		4.637		4.617	
Creatine:C2-H2	0.918	0.905	0.975	0.963	1.482	1.292	1.460	1.270
Creatine:CH3	0.893	1.93%	0.951	1.81%	1.101	20.83%	1.080	21.19%
Glu:Cγ-H2	10.215	11.257	10.261	11.302	11.524	11.982	11.507	11.965
Glu:Cβ-H2	12.298	13.09%	12.342	13.02%	12.440	5.41%	12.423	5.42%
Gly:Cα-H2	7.873	7.873	7.923	7.923	10.746	10.746	10.729	10.729
Ile:Cδ-H3	0.850	0.897	0.908	0.954	1.442	1.229	1.420	1.207
Ile:Cγ2-H3	0.943	7.29%	1.001	6.84%	1.015	24.55%	0.994	24.99%
Lactate:C2-H1	19.628	19.973	19.663	20.009	24.074	22.539	24.062	22.527
Lactate:CH3	20.319	2.45%	20.354	2.44%	21.003	9.63%	20.991	9.64%
Leu:Cα-H1	1.881	2.086	1.938	2.143	2.282	2.486	2.261	2.465
Leu:Cβ-H2	1.930	11.92%	1.986	11.59%	2.010	16.56%	1.989	16.71%

Compound: 2D NMR cross peak used in quantitation	Liver_extract with Fe(III)EDTA				Liver_extract without Fe(III)EDTA			
	Conc. Calcd. k,b (mM)	Ave./Precision	Conc. Calcd. k,b (mM)	Ave./Precision	Conc. Calcd. k,b (mM)	Ave./Precision	Conc. Calcd. k, 0 (mM)	Ave./Precision
Leu:Cδ2-H3 ¹¹	2.104		2.160		2.756		2.735	
Leu:Cδ1-H3	2.431		2.487		2.895		2.874	
Lys:Cβ-H2	1.323		1.380		2.203		2.182	
Lys:Cδ-H2	0.912	26.01%	0.969		2.614	24.71%	2.593	12.19%
Met:Cγ-H2*	0.437	0.432	0.495	0.490				
Met:Cε-H3	0.427	1.61%	0.485	1.41%	0.693		0.671	0.671
Ornithine:C3-H2	0.627	0.627	0.685	0.685	1.901		1.879	1.879
Pantothenate:CβH3*	0.181	0.368	0.240	0.426				
Pantothenate:CaH3* ¹¹	0.555	71.81%	0.613	61.90%				
Ser:Cα-H1	2.095	1.947	2.151	2.003	2.660		2.639	2.660
Ser:Cβ-H2	1.799	10.75%	1.856	10.44%	2.701	1.09%	2.680	1.10%
Succinate:(CH2)2	5.840	5.840	5.892	5.892	7.253		7.234	7.234
Taurine:C1-H2 ¹⁰	0.301	0.568	0.360	0.626	0.740		0.718	0.773
Taurine:C2-H2	0.835	66.42%	0.893	60.18%	0.849	9.65%	0.827	9.92%
Thr:Cα-H1	1.215	1.288	1.272	1.346	1.231		1.210	1.466
Thr:Cβ-H1	1.362	8.05%	1.419	7.70%	1.743	24.34%	1.722	24.71%
Val:Cα-H1	2.239	1.971	2.295	2.027	3.354		3.334	2.464
Val:Cγ1-H3	1.838	11.78%	1.895	11.44%	2.004	30.33%	1.983	30.61%
Val:Cγ2-H3	1.835		1.891		2.098		2.076	
α-Glucose:C1-H1	38.908	38.908	38.921	38.921	44.253		44.251	44.251
β-Glucose:C1-H1	71.314	71.314	71.288	71.288	75.910		75.922	75.922
%αG		35.30%		35.32%		36.83%		36.82%

Data were analyzed as in Table 1. ¹¹ and ¹⁰ signify overlapped peaks.

* Peaks (red) were missing in spectra of sample without Fe(III)EDTA.