Journal of Magnetic Resonance 347 (2023) 107365

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

# $^{13}$C NMR metabolomics: *J*-resolved STOCSY meets INADEQUATE

Mario Uchimiya [a], Malin Olofsson [b,c], McKenzie A. Powers [c], Brian M. Hopkinson [c], Mary Ann Moran [c], Arthur S. Edison [a,d,*]

[a] Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA
[b] Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Sweden
[c] Department of Marine Sciences, University of Georgia, Athens, GA 30602, USA
[d] Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

ARTICLE INFO

*Article history:*
Received 28 October 2022
Revised 20 December 2022
Accepted 28 December 2022
Available online 31 December 2022

*Keywords:*
Metabolomics
$^{13}$C
*J*-resolved
STOCSY
INADEQUATE
Phytoplankton

ABSTRACT

Robust annotation of metabolites is a challenging task in metabolomics. Among available applications, $^{13}$C NMR experiment INADEQUATE determines direct $^{13}$C–$^{13}$C connectivity unambiguously, offering indispensable information on molecular structure. Despite its great utility, it is not always practical to collect INADEQUATE data on every sample in a large metabolomics study because of its relatively long experiment time. Here, we propose an alternative approach that maintains the quality of information but saves experiment time. In this approach, individual samples in a study are first screened by $^{13}$C homonuclear *J*-resolved experiment (JRES). Next, JRES data are processed by statistical total correlation spectroscopy (STOCSY) to extract peaks that behave similarly among samples. Finally, INADEQUATE is collected on one internal pooled sample to select STOCSY peaks that originate from the same compound. We tested this concept using the $^{13}$C-labeled endometabolome of a model marine diatom strain incubated under various settings, intending to cover a range of metabolites produced under different external conditions. This scheme was able to extract known diatom metabolites proline, 2,3-dihydroxypropane-1-sulfonate (DHPS), β-1,3-glucan, choline, and glutamate. This pipeline also detected unknown compounds with structural information, which is valuable in metabolomics where *a priori* knowledge of metabolites is not always available. The ability of this scheme was seen even in sugar regions, which are usually challenging in $^{1}$H NMR due to severe peak overlap. JRES and INADEQUATE were highly complementary; INADEQUATE provided directly-bonded $^{13}$C networks, whereas JRES linked INADEQUATE networks within the same compound but broken by nitrogen or sulfur atoms, highlighting the advantage of this integrated approach.


## 1. Introduction

As metabolomics grows rapidly in diverse research areas, robust annotation of metabolites remains a challenging task, regardless of the platform [1,2]. In NMR metabolomics, various $^{1}$H-detected two-dimensional (2D) experiments that correlate $^{1}$H and $^{13}$C nuclei are the primary approaches in compound annotation [3,4]. Although $^{1}$H-based experiments have been successfully and widely implemented, $^{13}$C-detection experiments can complement $^{1}$H NMR and increase the confidence level of annotation [5]. $^{13}$C spectra have a much larger chemical shift dispersion than $^{1}$H, leading to less ambiguity arising from overlapping peaks, and $^{13}$C nuclei with no attached $^{1}$H can be detected, which can improve annotation. $^{13}$C chemical shifts are relatively stable to sample conditions such as differences in pH, temperature, and salt, facilitating robust use of reference spectra in databases. Most importantly from an annotation perspective, molecular backbone structure can be unambiguously determined based on direct $^{13}$C–$^{13}$C connectivity, fundamental information for structural elucidation [5].

Incredible natural abundance double-quantum experiment (INADEQUATE) is a $^{13}$C-detection experiment that correlates directly-bonded $^{13}$C nuclei [6]. This experiment creates double-quantum coherence associated with coupled spins, and the resulting spectrum elucidates molecular networks. Although INADEQUATE was originally proposed for natural abundance $^{13}$C, this experiment can benefit from isotope enrichment [7,8]. Because of a large $^{13}$C chemical shift dispersion along both the acquisition and double-quantum axes, INADEQUATE is especially valuable for complex mixtures that are typical for metabolomics samples [7]. Applying INADEQUATE to multiple samples in a study further

* Corresponding author at: Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Rd., Athens, GA 30602, USA.
*E-mail address:* aedison@uga.edu (A.S. Edison).

https://doi.org/10.1016/j.jmr.2022.107365
1090-7807/

enables multivariate analysis to extract compounds characterizing specific samples or groups [9]. Despite its great utility, it is often impractical to collect INADEQUATE data on every sample in a large study given that each experiment requires many hours of data collection. For broader application of this robust experiment to metabolomics, there is a need to obtain equivalent information with less experiment time. One obvious approach would be direct $^{13}$C 1D experiments on each sample, but for $^{13}$C-enriched samples, the resulting resonances are complicated and overlapped by numerous $^{13}$C–$^{13}$C couplings.

Here, we propose an alternative approach that maintains the quality of information but saves experiment time. In this approach, individual samples in a study are first profiled by $^{13}$C homonuclear *J*-resolved experiment (JRES), as opposed to INADEQUATE. $^{13}$C-JRES is a 2D $^{13}$C detection experiment that resolves chemical shifts and *J*-coupling constants along different dimensions. Importantly, the chemical shift acquisition dimension can be obtained with very high-resolution, and the *J*-coupling dimension requires relatively few points to resolve the $^{13}$C–$^{13}$C couplings, resulting in relatively short overall experiment times. The 2D $^{13}$C-JRES allows us to create projection spectra along the chemical shift dimension, effectively providing fully decoupled $^{13}$C spectra that retain high resolution. The $^{13}$C projection spectra can be analyzed with the multitude of statistical analysis tools that are routinely applied to standard $^{1}$H 1D NMR spectra such as PCA, PLS-DA, or OPLS-DA [10]. However, the $^{13}$C data have the added advantage of very little resonance overlap and require no or very little alignment. In this paper, we demonstrate the utility of $^{13}$C-JRES projection spectra by applying statistical total correlation spectroscopy (STOCSY) to investigate the correlation of intensities among all peaks across all the samples, an approach widely used in $^{1}$H-detection NMR [9]. Positive correlation between peaks in STOCSY indicates that those peaks behave similarly among samples, suggesting that they are potentially from the same compound. They, however, could also be from different compounds that are sharing the same biological pathway, or even compounds that are in different pathways but have the same biological response [9]. To complement the $^{13}$C-JRES data, we also collect a complete 2D INADEQUATE dataset on one internal pooled sample. The INADEQUATE and $^{13}$C-JRES data are highly complementary, because the INADEQUATE provides the directly-bonded $^{13}$C network, and $^{13}$C-JRES can link two or more $^{13}$C INADEQUATE networks within the same compound but broken by a heteroatom.

We tested this concept using the $^{13}$C-labeled endometabolome from a model marine phytoplankton strain, varying the incubation settings including temperature, nutrient condition, and the presence of co-culturing bacteria. Production of metabolites by phytoplankton in the ocean represents one of the key elements in global biogeochemical cycles [11]. Phytoplankton metabolite production can be variable depending on ambient conditions such as water temperature [12], nutrients [13], and the presence of bacteria [14]. This experimental design provides an ideal setting to test the concept.

## 2. Results and discussion

We acquired $^{13}$C-JRES spectra on phytoplankton endometabolome for all the samples in the study ($n$ = 24). We also obtained a single INADEQUATE spectrum using an internal pooled sample, a representative sample prepared by mixing aliquots from all the study samples. $^{13}$C-JRES projection spectra were analyzed using various STOCSY driver peaks, and this output was linked to the INADEQUATE spectrum for structural elucidation.

### 2.1. Linking $^{13}$C-JRES and INADEQUATE

We applied Bruker's standard JRES pulse program to an internal pooled sample using a Bruker NEO 600 MHz spectrometer with a 5-mm DCH probe (Fig. 1a, top). A skyline projection created from the JRES spectra yielded sharp singlet peaks (Fig. 1a, bottom), allowing us to apply downstream analysis without interference from extensive $^{13}$C–$^{13}$C *J* couplings.

We applied the same $^{13}$C-JRES experiment to all the 24 study samples, and they were analyzed by STOCSY. Fig. 1b shows a representative STOCSY spectrum showing correlations of peaks with a driver peak at 26.49 ppm. The number of peaks positively correlated (threshold, $r > 0.85$) with the driver peak was eight (26.49, 31.68, 48.81, 56.44, 63.95, 67.30, and 177.41 ppm; Fig. 1b). Those eight peaks could be from the same compound, but they could also be from multiple compounds from the same biological pathway, or even from different pathways with the same biological response. This ambiguity was clearly solved by introducing INADEQUATE (Fig. 1c). Out of those eight peaks, five of them were included in an INADEQUATE network that was from proline. The rest of the three peaks are statistically positively correlated peaks, but not from proline.

### 2.2. Generalization of STOCSY-INADEQUATE

We further applied this scheme to other driver peaks, intending to evaluate the ability of our approach to separate peaks and elucidate molecular structure for broader chemical space. For this evaluation, we first limited the number of peaks using a peak height threshold (0.03) and then looked for driver peaks that have at least one partner STOCSY peak. 64 driver peaks were extracted, and they covered a wide range of $^{13}$C chemical shifts (range, 18.9–177.4 ppm). Once STOCSY and INADEQUATE information was linked, corresponding compounds were searched for using a list of metabolites that was previously reported using the same diatom strain and NMR metabolomics (2D HSQC) [14].

In addition to proline, the test driver peaks also detected diatom compounds, including 2,3-dihydroxypropane-1-sulfonate (DHPS), β-1,3-glucan (main component of diatom laminarin), choline, and glutamate (Table 1, Supplementary Figure S1). In all the cases, this scheme was able to eliminate positively correlated peaks that were not originating from those compounds (Table 1). This highlights the effectiveness of INADEQUATE to reduce ambiguity caused by statistical correlations originating from other compounds. In our case, such additional correlations are from both known and unknown compounds.

This scheme also detected unknown compounds that were not on the reference list. For example, a driver peak at 79.47 ppm had positive correlation with 9 peaks based on STOCSY (Table 1). INADEQUATE revealed three of them are from a single compound that has a C—C—C backbone (Supplementary Figure S2, Unknown-A). Similarly, a driver peak at 74.78 ppm detected the presence of a compound with a C—C backbone (74.78 ppm-65.07 ppm, Unknown-B), at 62.94 ppm a compound with a C—C backbone (62.94 ppm-43.6 ppm, Unknown-C), and at 72.27 ppm a compound with a C—C backbone (72.27 ppm-78.66 ppm, Unknown-D). This STOCSY-INADEQUATE scheme is effective to elucidate molecular structure even if *a priori* knowledge on compounds is not available.

The ability of this scheme to separate peaks was seen even in complex sugar regions due to the broad $^{13}$C dispersion. As an example, the detected Unknown Compound-A (Table 1, Supplementary Figure S2a) is inside the network of β-1,3-glucan ($C_4$-$C_5$-$C_6$) (Table 1, Supplementary Fig. S1b). Even under this situation, JRES and STOCSY were able to separate those peaks. Also, INADEQUATE was able to provide structural information. In $^{1}$H NMR,

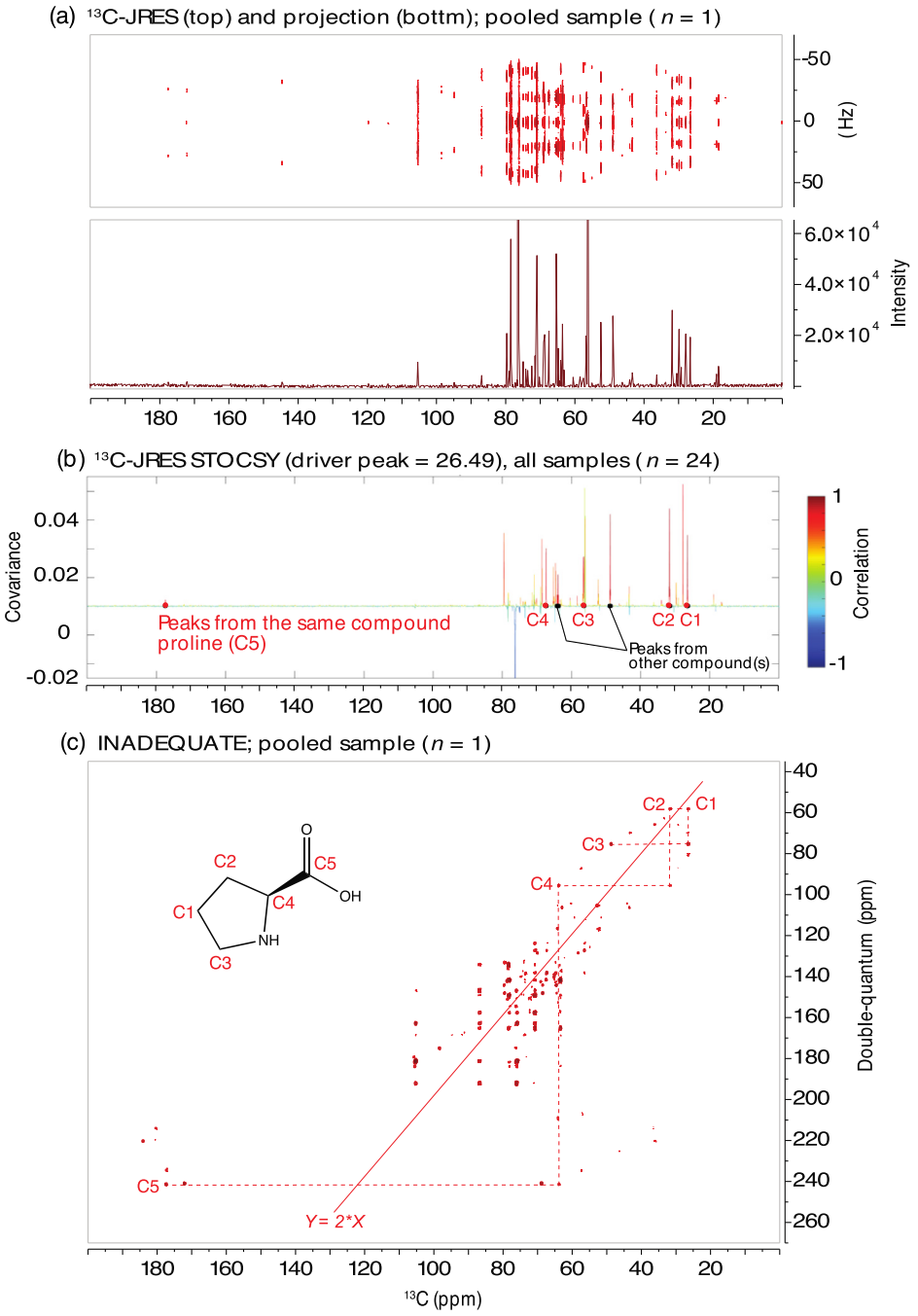


**Fig. 1.** (a) $^{13}C$ *J*-resolved (JRES) spectrum collected on a pooled internal standard sample (top) and its projection spectrum (bottom). (b) Example of statistical total correlation spectroscopy (STOCSY) spectrum based on JRES spectra collected on 24 study samples. The driver peak is 26.49 ppm. Circles on the baseline indicate peaks positively correlated ($r > 0.85$) with the driver peak. Among them, red circles are peaks originating from the same compound proline, whereas the black ones are peaks from other compound(s) (see below). (c) Incredible natural abundance double-quantum experiment (INADEQUATE) spectrum for the pooled sample. The solid red line depicts the double-quantum rule (Y = 2*X), and carbons next to each other have the same distance from this line. Dashed lines indicate carbon networks for the same compound proline. This information was used to distinguish peaks in Fig. 1b. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the corresponding regions for metabolomics samples are dominated by overlapping peaks and that makes the interpretation of STOCSY for this region very complicated.

During the test of driver peaks, we realized that some JRES peaks with high intensities were missing on INADEQUATE. We found that they were $^{13}C$ peaks isolated by nitrogen or sulfur in specific compounds, including betaine ($(CH_3)_3$-N-, 56.1 ppm) or dimethylsulfoniopropionate (DMSP; $(CH_3)_2$-S-, 27.8 ppm). Those $^{13}C$ resonances with no additional $^{13}C$ couplings are not detected by INADEQUATE but are detectable by JRES. This suggests that $^{13}C$-JRES and INADEQUATE should be used complementarily for metabolomics samples, rather than using them separately, for a broader coverage of compound information. Metabolites with isolated carbons are often seen and abundant in diatom samples in NMR metabolomics (e.g., DMSP, betaine, trimethylamine-oxide, and choline) [14,15].

As a result, out of the 64 driver peaks we tested, peaks that were linked to known compound information are 29, those for unknown

**Table 1**
Example of STOCSY-INADEQUATE outputs for test driver peaks. A complete list of peaks is also in Supplementary Table S1. Carbon identification numbers are based on Atom Label Assignment Tool using InChI String (ALATIS), except β-1,3-glucan for which conventional carbohydrate numbering was used. DHPS: 2,3-dihydroxypropane-1-sulfonate. *: Peaks with no significant STOCSY but observed in INADEQUATE. Note that the column "Peaks from the same compound" can be larger than the number of carbons in a compound because of the JRES projections, which in some cases can introduce multiple peaks from a single carbon.

| JRES STOCSY | | | | INADEQUATE | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Driver peak ppm | Correlated peaks | Peaks from the same compound | Cooccurring peak(s) | Carbon ppm | Double-quantum ppm (A-X) | Double-quantum ppm (X-A) | Compound | Carbon ID | Carbon network |
| 26.49 | 8 | 5 | 3 | 48.78 | 75.26 | – | Proline | 3 | C3-C1-C2-C4-C5 |
| | | | | 26.49 | 58.22 | 75.26 | | 1 | |
| | | | | 31.73 | 95.65 | 58.22 | | 2 | |
| | | | | 63.92 | 241.41 | 95.65 | | 4 | |
| | | | | 177.48 | – | 241.41 | | 5 | |
| 67.30 | 11 | 8 | 3 | 56.50 | 127.30 | – | DHPS | 2 | C2-C3-C1 |
| | | | | 70.80 | 138.10 | 127.30 | | 3 | |
| | | | | 67.30 | – | 138.10 | | 1 | |
| 78.34 | 8 | 6 | 2 | 105.4 | 181.40 | – | β-1,3-glucan | 1 | C1-C2-C3-C4-C5-C6 |
| | | | | 76.00 | 162.90 | 181.40 | | 2 | |
| | | | | 86.90 | 157.80 | 162.9 | | 3 | |
| | | | | 70.90 | 149.30 | 157.8 | | 4 | |
| | | | | 78.40 | 141.80 | 149.3 | | 5 | |
| | | | | 63.40 | – | 141.8 | | 6 | |
| 58.33 | 12 | 4 | 8 | 70.14 | 128.47 | – | Choline | 4 | C4-C5 |
| | | | | 58.33 | – | 128.47 | | 5 | |
| 36.20 | 4 | 2 | 2 | 184.09* | 220.29 | – | Glutamate | 4 | C4-C2-C1-C3-C5 |
| | | | | 36.20 | 65.93 | 220.29 | | 2 | |
| | | | | 29.73 | 87.09 | 65.93 | | 1 | |
| | | | | 57.36* | 234.72 | 87.09 | | 3 | |
| | | | | 177.36* | – | 234.72 | | 5 | |
| 79.47 | 9 | 3 | 6 | 68.51 | 148.10 | – | Unknown-A | – | C-C-C |
| | | | | 79.59 | 144.16 | 148.10 | | | |
| | | | | 64.57 | – | 144.16 | | | |
| 74.78 | 2 | 2 | 0 | 74.78 | 139.85 | – | Unknown-B | – | C-C |
| | | | | 65.07 | – | 139.85 | | | |
| 62.94 | 4 | 2 | 2 | 62.94 | 106.54 | – | Unknown-C | – | C-C |
| | | | | 43.60 | – | 106.54 | | | |
| 72.27 | 8 | 2 | 6 | 72.27 | 150.93 | – | Unknown-D | – | C-C |
| | | | | 78.66 | – | 150.93 | | | |

compounds are 11, and peaks that were not linked to INADEQUATE, partly due to the methodological difference in target carbons between JRES and INADEQUATE, was 24 (Supplementary Table S1). In addition to the known compounds listed in this study, the previous literature also reported other compounds (amino alcohol, glycerol derivative, nucleoside, and organic acids) for similar diatom samples based on HSQC [14], and this gap could partly be filled by decreasing a threshold to select driver peaks, enabling detection of lower concentration metabolites. For very low concentrations, the overall $^{13}C$ sensitivity will also lead to fewer compounds compared to $^{1}H$-detected experiments.

### 2.3. Applicability of JRES-INADEQUATE to broader metabolomics studies

Although specialized custom $^{13}C$ probes can improve $^{13}C$-detected experiments [16–19], we conducted this study using a commercial Bruker DCH cryoprobe and default pulse programs. The scheme proposed here can be broadly adopted by other laboratories. It should be noted that the JRES pulse program we used in this study (jresdcqf) does not use an adiabatic 180-degree pulse. We verified that non-adiabatic and adiabatic 180-degree pulses did not make a noticeable difference on JRES spectra on our 14.1 T magnet (600 MHz system; Supplementary Figure S3). On the other hand, an adiabatic pulse was critical when we tested the $^{13}C$-JRES experiment at 21.1 T (900 MHz; Supplementary Figure S3). The adiabatic version of $^{13}C$-JRES is highly recommended at high magnetic fields.

### 2.4. Making the best use of $^{13}C$ NMR in metabolomics

Although the proposed scheme can reduce experiment time, the integration of STOCSY and INADEQUATE information is currently labor-intensive. They include linking STOCSY and INADEQUATE, creating INADEQUATE networks, and compound identification. Metabolomics deals with complex mixtures of compounds, and automation of this step is highly desired from a throughput perspective.

Due to the advantages of $^{13}C$ NMR, the scheme introduced here is ideal for computational tasks. INADEQUATE has a robust double-quantum rule ($Y = 2^*X$) that constrains neighboring carbons and networks. $^{13}C$ NMR chemical shift values are relatively tolerant to sample conditions and suitable for database query. INADEQUATE reference spectra are not widely available in databases, but accurate simulated INADEQUATE networks can be created using compound structure with assigned $^{13}C$ chemical shifts from databases such as BMRB [20]. Experimental INADEQUATE networks can then be semi-automatically detected and matched to the simulated database, as previously described by our group. We are currently updating our INADEQUATE network analysis

algorithm (INETA) and database [7], which was originally implemented using Mathematica, to Python; this will be reported elsewhere. This will enable us to query unknown networks found in this study (C—C and C—C—C) using large databases, providing a path to annotate unknown compounds. Even when no database match is found for those unknown compounds, INADEQUATE data can be further used for *de novo* compound identification, especially when mass spectrometry and other 2D NMR data are available using established approaches widely used in natural products chemistry [21]. Other NMR experiments such as ADEQUATE [22] and TOCSY [23] could, in principle, be used in place of INADEQUATE in our scheme, but these typically employ $^{1}H$ detection that limits the resolution in the acquisition dimension. TOCSY can also be used with $^{13}C$ detection [24], but the advantage of INADEQUATE is its ability to provide definitive carbon connectivity.

## 3. Material and methods

### 3.1. Sample preparation

Six treatments of a marine diatom strain *Thalassiosira pseudonana* CCMP1335 were prepared: treatments incubated axenically at either 14, 20, or 28 °C, and treatments co-cultured with a bacterial strain *Ruegeria pomeroyi* DSS-3 at the corresponding temperatures (four replicates for each). L1 media [25] was used with $NaH^{13}CO_3$ as a source of bicarbonate. The diatom used for the co-cultured treatments was $B_{12}$ stressed to emphasize the known co-existing system [13]. The light cycle consisted of 16 h light (120 μmol photons $m^{-2}$ $s^{-1}$) and 8 h of dark.

Diatom cells were filtered onto 2.0-μm pore size 47-mm diameter filters (Isopore) and kept in −80 °C until further processing. Filters were transferred to Milli-Q water (Millipore) in 50-mL tubes, and cells were removed from filters by sonication [26]. Samples were moved to a −80 °C freezer, and then freeze-dried using a lyophilizer (LABCONCO). The dried samples were added to a deuterated phosphate buffer (30 mmol/L, pH 7.4) and an internal standard of 2,2-dimethyl-2-silapentane-5-sulfonate-$d_6$ (DSS, 1 mmol/L), vortexed for 5 min, centrifuged at 20,800 *rcf* for 10 min using an ultracentrifuge (Eppendorf), and supernatants were transferred to 5-mm NMR tubes (NORELL). Throughout the sample processing, samples were kept on ice or at 4 °C.

### 3.2. NMR data acquisition, processing, and analysis

NMR experiments were conducted using an Avance Neo 600 MHz NMR spectrometer equipped with a 5-mm DCH cryoprobe (Bruker) at 298 K. Bruker's default pulse programs were used for the following carbon detection experiments: 'jresdcqf' for carbon homonuclear *J*-resolved correlation experiment (JRES) with proton decoupling; 'inadphppsp' for incredible natural abundance double-quantum experiment (INADEQUATE). Parameter settings used for those experiments are described in Supplementary Table S2. Experiments were conducted using TopSpin version 4.0.6 (Bruker).

The raw Bruker spectra data were processed by MestReNova version 14.3 (Mestrelab Research). Detailed processing parameters are in Supplementary Table S3 and are also available as deposited files (see below). Briefly, for JRES, data were subjected to apodization with sine bell ll for both f1 and f2 dimensions. Spectra were further symmetrized, tilted, and referenced to DSS. They were further converted to skyline projections and baseline-corrected using polynomial fit. Those output spectra were saved as a matrix (.csv) and used in the downstream analysis. For INADEQUATE, data were processed using apodization with sine for both f1 and f2 dimensions.

For STOCSY analysis, processed JRES projection spectra were loaded and analyzed using MATLAB version R2022a (MathWorks) and metabolomics toolbox (https://github.com/artedison/Edison_Lab_Shared_Metabolomics_UGA). Projection spectra were aligned (function 'guide_align1D') using PAFFT with a mean distance metric, normalized with probable quotient normalization (PQN) [27] ('normalize'), and subjected to STOCSY analysis ('STOCSY'). For atom numbering convention, Atom Label Assignment Tool using InChI String (ALATIS) was used [28].

All the Bruker time-domain data, MestReNova processing files and processed files, and MATLAB scripts were deposited to Metabolomics Workbench with Study ID ST00232.

## CRediT authorship contribution statement

**Mario Uchimiya:** Writing – original draft, Conceptualization, Investigation, Formal analysis, Visualization. **Malin Olofsson:** Writing – review & editing, Resources. **McKenzie A. Powers:** Writing – review & editing, Resources. **Brian M. Hopkinson:** Writing – review & editing, Supervision, Funding acquisition. **Mary Ann Moran:** Writing – review & editing, Supervision, Funding acquisition. **Arthur S. Edison:** Writing – original draft, Conceptualization, Formal analysis, Visualization, Supervision, Funding acquisition.

## Data availability

Original data, processing scripts, and processed files are available at Metabolomics Workbench with Study Identification ST002321 (DOI: http://dx.doi.org/10.21228/M8HM7D).

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The authors declare no competing interests. We thank John Glushka for NMR expertise, Frank Ferrer-González and Jeremy Schreier for experimental help, and Rahil Taujale and Ricardo Borges for useful discussions. This work was supported by NSF (grant numbers 1948104 awarded to B.M.H., A.S.E, and M.A.M, OCE-2019589 awarded to A.S.E and M.A.M, and 1946970 awarded to A.S.E) and NIH (5R01GM120151-04 to A.S.E.). Original data, processing scripts, and processed files are available at Metabolomics Workbench with Study Identification ST002321 (DOI: http://dx.doi.org/10.21228/M8HM7D). The experimental workflows and protocols are being developed as part of the NSF Network for Advanced NMR (NAN) (1946970). We thank the reviewers for their productive comments. This is the NSF Center for Chemical Currencies of a Microbial Planet (C-CoMP) publication #017.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmr.2022.107365.

## References

[1] M.P.M. Letertre, G. Dervilly, P. Giraudeau, Combined nuclear magnetic resonance spectroscopy and mass spectrometry approaches for metabolomics, Anal. Chem. 93 (1) (2021) 500–518.
[2] M.E. Monge, J.N. Dodds, E.S. Baker, A.S. Edison, F.M. Fernandez, Challenges in identifying the dark molecules of life, Annu. Rev. Anal. Chem. 12 (12) (2019) 177–199.

[3] K. Bingol, R. Bruschweiler, Multidimensional approaches to NMR-based metabolomics, Anal. Chem. 86 (1) (2014) 47–57.
[4] K. Bingol, D.W. Li, B. Zhang, R. Bruschweiler, Comprehensive metabolite identification strategy using multiple two-dimensional NMR spectra of a complex mixture implemented in the COLMARm web server, Anal. Chem. 88 (24) (2016) 12411–12418.
[5] A.S. Edison, A. Le Guennec, F. Delaglio, E. Kupce, Practical guidelines for $^{13}$C-based NMR metabolomics, Methods Mol. Biol. 2037 (2019) 69–95.
[6] A. Bax, R. Freeman, S.P. Kempsell, Natural abundance $^{13}$C–$^{13}$C coupling observed via double-quantum coherence, J. Am. Chem. Soc. 102 (14) (1980) 4849–4851.
[7] C.S. Clendinen, C. Pasquel, R. Ajredini, A.S. Edison, $^{13}$C NMR metabolomics: INADEQUATE network analysis, Anal. Chem. 87 (11) (2015) 5698–5706.
[8] B.H. Oh, W.M. Westler, P. Darba, J.L. Markley, Protein carbon-13 spin systems by a single two-dimensional nuclear magnetic resonance experiment, Science 240 (4854) (1988) 908–911.
[9] O. Cloarec, M.E. Dumas, A. Craig, R.H. Barton, J. Trygg, J. Hudson, et al., Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic $^{1}$H NMR data sets, Anal. Chem. 77 (5) (2005) 1282–1289.
[10] J. Debik, M. Sangermani, F. Wang, T.S. Madssen, G.F. Giskeodegard, Multivariate analysis of NMR-based metabolomic data, NMR Biomed. 35 (2) (2022) e4638.
[11] M.A. Moran, E.B. Kujawinski, W.F. Schroer, S.A. Amin, N.R. Bates, E.M. Bertrand, et al., Microbial metabolites in the marine carbon cycle, Nat. Microbiol. 7 (4) (2022) 508–523.
[12] A. Toseland, S.J. Daines, J.R. Clark, A. Kirkham, J. Strauss, C. Uhlig, et al., The impact of temperature on marine phytoplankton resource allocation and metabolism, Nat. Clim. Chang. 3 (11) (2013) 979–984.
[13] B.P. Durham, S. Sharma, H.W. Luo, C.B. Smith, S.A. Amin, S.J. Bender, et al., Cryptic carbon and sulfur cycling between surface ocean plankton, Proc. Nat. Acad. Sci. 112 (2) (2015) 453–457.
[14] M. Uchimiya, W. Schroer, M. Olofsson, A.S. Edison, M.A. Moran, Diel investments in metabolite production and consumption in a model microbial system, ISME J. 16 (5) (2022) 1306–1317.
[15] M. Olofsson, F.X. Ferrer-González, M. Uchimiya, J.E. Schreier, N.R. Holderman, C.B. Smith, et al., Growth-stage-related shifts in diatom endometabolome composition set the stage for bacterial heterotrophy. ISME, Communications 2 (28) (2022), https://doi.org/10.1038/s43705-022-00116-5.
[16] O. Sanati, A.S. Edison, L.A. Hornak, I.M. Litvak, V. Ramaswamy, N. Freytag, et al., $^{13}$C-optimized HTS NMR RF coil design at 21.1 T, IEEE Trans. Appl. Supercond. 31 (5) (2021) 1–5.
[17] J.N. Thomas, V. Ramaswamy, I.M. Litvak, T.L. Johnston, A.S. Edison, W.W. Brey, Progress towards a higher sensitivity $^{13}$C-optimized 1.5 mm HTS NMR probe, IEEE Trans. Appl. Supercond. 31 (5) (2021) 1–4.
[18] V. Ramaswamy, J.W. Hooker, R.S. Withers, R.E. Nast, A.S. Edison, W.W. Brey, Microsample cryogenic probes: Technology and applications, eMagRes 2 (2013) 215–228.
[19] V. Ramaswamy, J.W. Hooker, R.S. Withers, R.E. Nast, W.W. Brey, A.S. Edison, Development of a $^{13}$C-optimized 1.5-mm high temperature superconducting NMR probe, J. Magn. Reson. 235 (2013) 58–65.
[20] E.L. Ulrich, H. Akutsu, J.F. Doreleijers, Y. Harano, Y.E. Ioannidis, J. Lin, et al., BioMagResBank, Nucleic Acids Res. 36 (2008) D402–D408.
[21] A.S. Edison, F.C. Schroeder, NMR - Small molecules and analysis of complex mixtures, in: H.W. Liu, L. Mander (Eds.), Modern Methods in Natural Product Chemistry, Elsevier, Oxford., 2010, pp. 169–196.
[22] J. Weigelt, G. Otting, $^{1}$H-detected INEPT-INADEQUATE at natural $^{13}$C abundance, J. Magn. Reson. A 113 (1) (1995) 128–130.
[23] A. Bax, G.M. Clore, A.M. Gronenborn, $^{1}$H–$^{1}$H correlation via isotropic mixing of $^{13}$C magnetization, a new three-dimensional approach for assigning $^{1}$H and $^{13}$C spectra of $^{13}$C-enriched proteins, J. Magn. Reson. 88 (2) (1990) 425–431.
[24] A. Eletsky, O. Moreira, H. Kovacs, K. Pervushin, A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using $^{13}$C spectroscopy, J. Biomol. NMR 26 (2) (2003) 167–179.
[25] R.R.L. Guillard, P.E. Hargraves, *Stichochrysis immobilis* is a diatom, not a chyrsophyte, Phycologia 32 (3) (1993) 234–236.
[26] M. Uchimiya, Y. Tsuboi, K. Ito, Y. Date, J. Kikuchi, Bacterial substrate transformation tracked by stable-isotope-guided NMR metabolomics: application in a natural aquatic microbial community, Metabolites 7 (4) (2017) 52, https://doi.org/10.3390/metabo7040052.
[27] F. Dieterle, A. Ross, G. Schlotterbeck, H. Senn, Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in H-1 NMR metabonomics, Anal. Chem. 78 (13) (2006) 4281–4290.
[28] H. Dashti, W.M. Westler, J.L. Markley, H.R. Eghbalnia, Unique identifiers for small molecules enable rigorous labeling of their atoms, Sci. Data 4 (2017).