Combined Analysis of NMR and MS Spectra (CANMS)

Ludwig, Christian; Chong, Mei G; Jayaraman, Anusha; Marin, Silvia; Selivanov, Vitaly; de Atauri Carulla, Pedro; Tennant, Daniel; Cascante, Marta; Gunther, Ulrich

DOI:
10.1002/anie.201611634

Document Version
Peer reviewed version

Citation for published version (Harvard):
Ludwig, C, Chong, MG, Jayaraman, A, Marin, S, Selivanov, V, de Atauri Carulla, P, Tennant, D, Cascante, M & Gunther, U 2017, 'Combined Analysis of NMR and MS Spectra (CANMS)', Angewandte Chemie (International Edition) , vol. 56, no. 15, pp. 4140-4144. https://doi.org/10.1002/anie.201611634

Combined NMR and MS analysis for tracer based metabolic flux experiments

Ludwig, Christian; Chong, Mei G; Jayaraman, Anusha; Marin, Silvia; Selivanov, Vitaly; de Atauri Carulla, Pedro; Tennant, Daniel; Cascante, Marta; Gunther, Ulrich

Combined NMR and MS analysis for tracer based metabolic flux experiments

Mei Chong[b], Anusha Jayaraman[c], Silvia Marin[c], Vitaly Selivanov[c], Pedro R. de Atauri Carulla[c], Daniel A. Tennant[d], Marta Cascante[c], Ulrich L. Günther[b], Christian Ludwig*[a]

Abstract: Cellular metabolism in mammalian cells represents an emerging challenge for analytical chemistry in the context of current biomedical research. Mass spectrometry (MS) and Nuclear Magnetic Resonance (NMR) spectroscopy in connection with computational tools have been used to study metabolism in cells. Compartmentalization of metabolism complicates the interpretation of stable isotope patterns in mammalian cells owing to the superimposition of different pathways contributing to the same pool of analytes in whole cell extracts. This indicates a need for a model-free approach to interpret such tracer-based data. MS and NMR provide complementary analytical information for metabolites with partial label incorporation in different positions. Here we present an approach that first simulates ^{13}C-multiplets in NMR spectra and utilizes mass increments to obtain long-range information. In a second step the combined information is utilized to derive isotopomer distributions for key metabolites. This is a first rigorous analytical and computational approach for a model-free analysis of metabolic flux data applicable to mammalian cells.

With every change in the phenotype of a cell or tissue, a supportive or compensatory change in metabolism is required. It therefore comes as no surprise that most, if not all, diseases are associated with, or in some cases driven by, changes in metabolism[1–3]. However, metabolic networks in mammalian cells are complex[4–7]. Despite existing models[8,9] our understanding of the mammalian metabolic network is still incomplete. Specific metabolic pathways can be studied using tracer-based approaches after exposing cells or tissues to a metabolic precursor enriched with one or more stable isotopes[10–12] (e.g. ^{13}C-labelled glucose or glutamine). When the enriched metabolic precursors are metabolized label is incorporated into various metabolites, giving rise to a potentially large number of isotopomers[13–16]. The isotopomer distributions depend on the metabolic pathways through which the labelled precursors have been metabolized. For instance, for cells grown in the presence of [1,2-$^{13}C_2$]-D-glucose an isotopomer analysis for C4 of glutamate can be used to infer flux through pyruvate dehydrogenase (PDH) into the TCA cycle (also known as the Krebs cycle). ^{13}C-observed NMR spectra provide information on isotopomer compositions, as multiplet structures alter with ^{13}C label incorporation in metabolites[17,18]. This is however limited to large concentrations of metabolites even when ^{13}C optimized NMR probes are used. ^{1}H-^{13}C-Heteronuclear Single Quantum Coherence (HSQC) spectra are substantially more sensitive although limited to proton bearing carbons., For this HSQC spectra need to be collected with a high number of increments to observe J_{CC} couplings, which is now best achieved using non-uniformly samples spectra (NUS). For example, much of the activity of the TCA cycle can be derived from multiplet patterns in molecules such as glutamate which reflect the mechanisms how the molecule was produced. This can either be through entry of pyruvate into the TCA via pyruvate dehydrogenase (PDH), or alternatively via pyruvate carboxylase (PC), but can also be a transamination product of glutamine or protein degradation.

Additional information on isotopomers can be derived from mass spectrometry[19,20], which yields mass increments rather than atom specific information. This provides crucial information in conjunction with the aforementioned NMR analysis which is complementary to the site-specific NMR information[21,22]. Here we present an algorithm for a bias-free determination of isotopomer compositions. The approach consists of two steps. The first step is to simulate the NMR spectrum for the different isotopomers. An NMR signal is only observed if at least one position is ^{13}C-labelled; in MS-terms this means that an m+1 or higher-labelled molecule is observed.

Isotopomer analysis is split into two parts as shown in Figure 1. First NMR multiplets are simulated using the pyGamma[23] library. Because the HSQC NMR spectra were acquired using echo/anti-echo coherence selection, development of carbon-carbon couplings is not zero in the first increment. To minimise computational time, the effect of the non-zero J_{CC} contribution is mimicked by the simulation of a ^{13}C spin-echo before the simulation of the ^{13}C free induction decay. This procedure produces good results for molecules in the absence of strongly coupled carbon-carbon spin systems, an assumption that is valid for all multiplets derived from ^{13}C-labelled glucose. The result of this simulation is shown in Figure 2 for glutamate C2 in MCF7 cells that were labelled with [1,2-^{13}C] glucose. Multiplets for other metabolites are shown in Figures S1-S5. This analysis simulates the sum of sub spectra arising from the label incorporation in adjacent carbons. For example, the simulation of glutamate C2 simulates contributions of spectra arising from the J_{C2C1}, J_{C2C3}, J_{C2C1C3} couplings and of the uncoupled spectrum. This analysis yields *local* ^{13}C-labelling pattern for C2, disregarding the 12C-contribution of C2 which is not observed in NMR spectra as well as remote ^{13}C-incorporations. The main focus of this algorithm is on the relative multiplet contributions which is independent of the polarisation transfer function in HSQC spectra and therefore independent of the JCH coupling constant.

In a second step the theoretical MS isotopologues, the HSQC multiplet percentages for all NMR visible carbon nuclei and the absolute ^{13}C enrichment percentages per carbon (1D NMR data)

[a] C. Ludwig
Institute of Metabolism and Systems Research
University of Birmingham
IBR West Tower, Birmingham, B15 2TT
E-mail: C.Ludwig@bham.ac.uk
[b] M. Chong, U.L. Günther, Institute of Cancer and Genome Sciences, University of Birmingham, Birmingham, UK
[c] A. Jayaraman, S. Marin, V. Selivanov, PR. De Atauri Carulla, M. Cascante, Department of Biochemistry and Molecular Biology, Faculty of Biology, Universitat de Barcelona, Spain
D. Tennant, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK
[d] D. Tennant, Institute of Metabolism and Systems Research, University of Birmingham, Birmingham, UK

Supporting information for this article is given via a link at the end of the document.

are simulated using a proposed isotopomer distribution (Fig. S6). This simulated NMR/MS data is then fitted to the experimental values using a least squares minimisation procedure by varying the proposed isotopomer distribution percentages. While at the beginning of the fitting procedure (without choosing which isotopomers are most likely present), the 1D NMR data are quite important, because the addition of per carbon ^{13}C percentages guides the fitting algorithm by reducing the amount of local minima; at latter stages of the fitting procedure (i.e. after the final set of isotopomers has been chosen based on all the data), it is beneficial to finalise the fit using just the NMR multiplet percentages and the MS isotopologue data as the 1D NMR data because of spectral congestion.

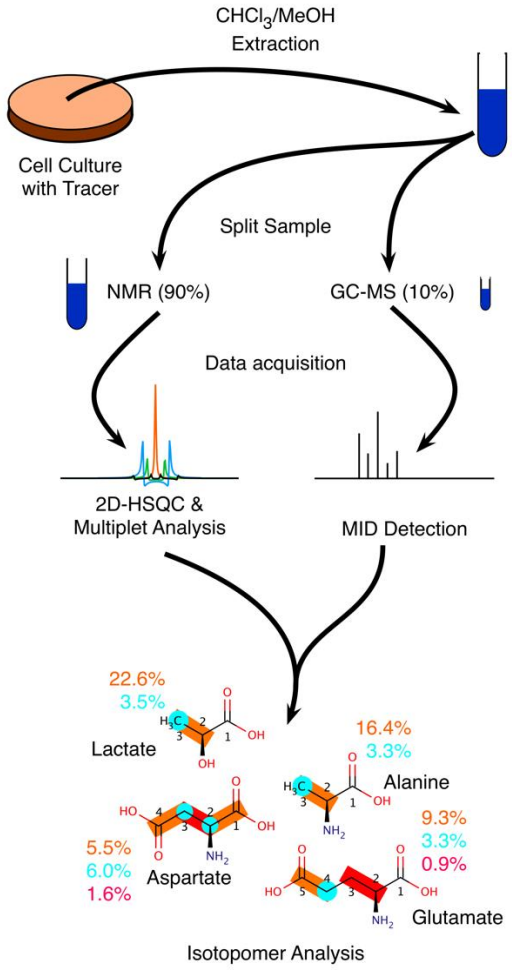

Figure 1. Schematic representation of model-free isotopomer analysis. After samples have been labelled using ^{13}C-enriched metabolic precursors, the cells are harvested and extracted using a modified Bligh-Dyer procedure. Each sample is then split before drying, where 90% are further prepared for NMR analysis and 10% for MS. 2D-^{1}H,^{13}C-HSQC spectra are acquired, mass spectra are acquired after derivatisation. In a first step, HSQC signals are simulated considering contributions of the possible J_{CC}-couplings contributing to the multiplet (c.f. Fig. 2). In a second step, the isotopomer composition is optimised for each metabolite to fit the relative composition of each multiplet together with the MIDs from the mass spectra. At no point during this analysis any assumptions are made about metabolic activity or pathways.

The analysis of metabolism in mammalian cells is complicated by the fact that several metabolic mechanisms contribute to the same metabolite pools. This is exacerbated in tracer-based experiments where isotopomers arise from label incorporation in different processes. In prokaryotic systems[24,25] computational metabolic flux is usually sufficient to quantify contributions from individual pathways. In eukaryotic cells mechanisms are more complex and therefore a model-free approach is desirable. Here we show that a model-free approach that simultaneously analyses signal multiplets in ^{1}H-^{13}C-HSQC NMR spectra and MIDs derived from GC-MS spectra, without assuming a metabolic model, yields site-specific isotopomer information for a range of key metabolites.

The process of this analysis is depicted in Fig. 1. The analysis is carried out in two steps, first by simulating the relative contributions of different coupling patterns to each multiplet. It is important to recognize that NMR by nature observes m+1 and above isotopomers only, because the ^{12}C-nucleus is not observable. Hence relative label incorporations in relation to the observed atom are obtained in this step. In the second step, this NMR information is combined between different atoms of one molecule and with the MIDs from MS. At no point during this analysis any assumptions are made about metabolic activity or pathways. Therefore, this approach represents a model-free isotopomer analysis of metabolic tracer data.

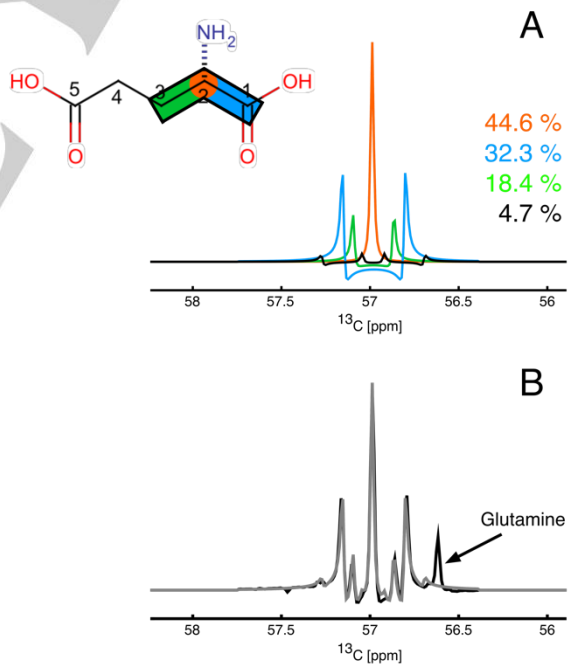

Figure 2. (**A**) Multiplet analysis of glutamate C2. Panel A depicts the deconvolution of the different constituting parts of the C2 multiplet of glutamate. The multiplet basically consists of 4 different components: glutamate with (i) C2 labelled (amber), (ii) C1 and C2 labelled simultaneously (blue), (iii) C2 and C3 labelled simultaneously (green) and (iv) C2, C1 and C3 labelled simultaneously (black). **(B)** shows the experimental spectrum of glutamate C2 in black and the fitted multiplet in grey. Overlap with the C2 or glutamine does not impair the analysis. (See supporting information for more detailed information on the implementation of multiplet analysis)

Fig. 2 shows an exemplary analysis for glutamate C2 where contributions of J_{C1C2}, J_{C1C2C3}, J_{C2C3} are considered in addition to C2 only labelled glutamate. $^{2}J_{CC}$ couplings are not resolved and therefore don't contribute to distinguishable signals in this spectrum. Therefore, the information extracted from this type of analysis is local in nature. As seen in Fig. 2B (and in Figs S1-S5) good simulations can be obtained for the signals of several metabolites, including lactate, alanine, glutamate and aspartate. These are key reporters for central metabolism. Fig 3. Illustrates the quality of fit obtained for several metabolites in the MCF7 cells used for this analysis. The analysis of a model mixture composed of [1-^{13}C], [1,2-^{13}C] and [U-^{13}C] glucose is shown in Fig. S7.

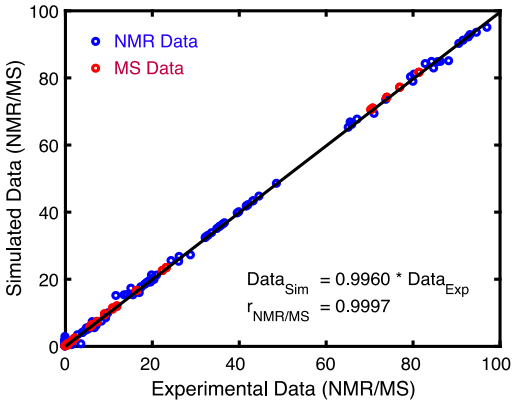

Figure 3. Quality of data analysis. Percentage contribution of multiplets and MIDs are reported for a simulation of the isotopomers of pyruvate, lactate, alanine, glutamate and aspartate. Data was used from 3 replicate samples of labelled MCF7 cells.

Considering the orthogonal nature of the two analytical techniques used in this approach, the question arises whether the two independent data sets lead to the same quantitative conclusions. Fig. 3 demonstrates that the NMR & MS data sets are highly consistent with each other, as supported by virtually perfect agreement between simulated and experimental data with a correlation coefficient of 99.97%. The combination of 2D-HSQC NMR and MS data together with [1,2-^{13}C] or [U-^{13}C] is unique because of the 1.07% natural abundance of ^{13}C. Since most of the singly labelled (i.e. no J coupling contributions) signals contain considerable contributions from unlabelled material, the area of highest accuracy for 2D-HSQC NMR data is between 1 and 10 % label incorporation. This is in contrast to other homonuclear NMR approaches such as 2D-^{1}H,^{1}H-TOCSY spectra or MS data which have their lowest accuracy between 1 and 10%. It is the combination of these two complementary techniques that enables a precise and accurate determination over the entire potential labelling percentages range.

This type of analysis allows for a series of important conclusions. For example, as shown in Fig. 4, lactate and alanine show characteristic labelling patterns in cells labelled with [1,2-$^{13}C_2$] glucose. If entirely produced by glycolysis C2 and C3 will always be jointly labelled and should hence show equal label incorporation. However, in the presence of pentose phosphate pathway activity, where the C1 of glucose is removed in the first step, label incorporation will be different in these positions[26,27]. In this analysis, we observe approx. 3.5% PPP contribution observed for lactate and pyruvate, whereas glycolysis contributes 22.5-23.5% to the label incorporation. Given that the glucose molecule is split into two parts, with one unlabelled for [1,2-^{13}C] glucose, this is equivalent to 7% PPP, and 45% glycolytic activity, respectively. Results from the isotopomer analysis are depicted in Fig. 1 for lactate, alanine, glutamate and aspartate.

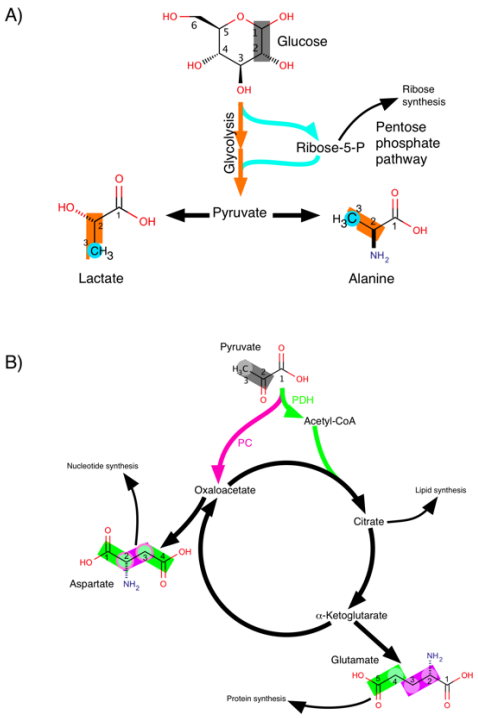

Figure 4. Tracing central metabolism using [1,2-^{13}C] glucose. Panel A: Glucose is primarily converted to pyruvate, either via glycolysis or via the pentose phosphate pathway (including both, the oxidative and the non-oxidative part). Depending on the pathway, either [2,3-^{13}C] or [3-^{13}C] pyruvate is produced. Read-out surrogates for glycolysis vs PPP usage are lactate, which is exclusively produced in the cytosol of mammalian cells and alanine, which can be produced both in the cytosol as well as mitochondria. Panel B: Pyruvate can then enter the citric acid cycle via two entry points. It can either be converted to acetyl-CoA by pyruvate dehydrogenase and then be condensated with oxaloacetate to form citrate, or it can be carboxylated to form oxaloacetate which then can be transformed into citrate or other molecules such as aspartate. Again, the labelling patterns in reporter molecules such as glutamate or aspartate are depending on the metabolic pathways used to produce these molecules.

Similarly, the entry point for pyruvate into the TCA cycle, through PDH or alternatively through the anaplerotic PC, generates different isotopomers (in Fig. 4B magenta for PC, and green for PDH). Isotopomers of TCA read-out molecules such as glutamate, malate and aspartate will inform on the presence and relative amount of PC activity. [4,5-$^{13}C_2$] glutamate is produced through the activity of PDH on pyruvate, while [2,3-$^{13}C_2$] glutamate is produced by PC activity. The [1,2-$^{13}C_2$] and [3,4-$^{13}C_2$] isotopomers of aspartate are also observed as a result of PDH,

while PC activity produces [2,3-$^{13}C_2$] aspartate. This type of analysis has recently been used in three reports to identify relative contributions of different metabolic pathways[28–30].

In the future, this approach may provide an avenue to identify contributions of metabolism from different cellular compartments. In a simple one-compartment model one would expect equal labelling patterns for metabolites such as lactate and alanine derived from pyruvate as a common precursor. Label incorporation in these metabolites, further down-stream, should also be reflected by labelling in glutamate and aspartate. In our model-free analysis of label incorporation in MCF7 cells, we observe equal label incorporation ratios between C2 and C3 for pyruvate and lactate, whereas glycolytically derived alanine is different. This effect is more pronounced for glutamate isotopomers. It is well-known that lactate is produced in the cytosol, whereas glutamate with this labelling pattern is exclusively synthesized in mitochondria. Alanine, on the other hand, can be synthesized from pyruvate in the cytosol as well as in mitochondria. These results can potentially be interpreted as alanine being preferentially synthesized inside mitochondria using predominantly PPP derived pyruvate. In combination with genetically modified cells (e.g. mitochondrial pyruvate carrier knock-outs) it should be possible to decipher such pathway contributions. The tools described here, which the authors distribute via NMRLab and MetaboLab[31,32], are uniquely qualified to describe metabolic pathways contributions in mammalian cells in a quantitative manner.

In conclusion, we present a model-free metabolic analysis incorporating quantitative NMR and MS data to obtain isotopomer contributions in ^{13}C-labelled breast cancer cells. Joint NMR/MS analysis leads to highly accurate and precise data that can be interpreted in the context of metabolic pathway analysis.

Experimental Section

MCF7 breast cancer cells were grown in the presence of [1,2-$^{13}C_2$]-D-glucose (Sigma-Aldrich) for 8h and subsequently extracted using a modified Blye-Digher extraction (see Supplementary Material for details). For NMR, samples were dissolved in 100mM phosphate buffer at pH 7.0.

The NMR analysis was carried out using ^{1}H-^{13}C-HSQC spectra acquired at 600MHz with 8192 complex increments to resolve ^{13}C-multiplets. In order to reduce acquisition times only 2048 data points were sampled non-uniformly. NMR spectra were reconstructed using compressed sensing[33] and NMRpipe[34] software.

For GC-MS, samples were run using a 7890A gas chromatograph coupled with an Agilent 5975C mass spectrometer and a HP-5 capillary column.

More details are presented in the supporting information.

Acknowledgements

Mei Chong and Anusha Jayaraman were supported by the EU grant METAFLUX FP7-PEOPLE-ITN-2010-264780. Ulrich Günther and Marta Cascante acknowledge support from the European Commission (Metaflux FP7-PEOPLE-ITN-2010-264780 and Cosmos EC- 312941). Marta Cascante also acknowledges the Spanish Ministerio de Economia y Competitividad (SAF2014-56059-R) and the Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) Generalitat de Catalunya (2014SGR1017) and the support received through the prize "ICREA Academia" for excellence in research, funded by ICREA foundation-Generalitat de Catalunya. We also thank TJ Ragan for helping with NUS data processing. The authors thank the Wellcome Trust for supporting the HWB-NMR facility in Birmingham.

Keywords: Metabolism, tracer, NMR, MS

[1] E. Holmes, I. D. Wilson, J. K. Nicholson, *Cell* **2008**, *134*, 714–717.
[2] C. M. Metallo, M. G. Vander Heiden, *Mol. Cell* **2013**, *49*, 388–398.
[3] R. J. DeBerardinis, C. B. Thompson, *Cell* **2012**, *148*, 1132–1144.
[4] P. S. Ward, C. Lu, J. R. Cross, O. Abdel-Wahab, R. L. Levine, G. K. Schwartz, C. B. Thompson, *J. Biol. Chem.* **2013**, *288*, 3804–3815.
[5] C. A. Lewis, S. J. Parker, B. P. Fiske, D. McCloskey, D. Y. Gui, C. R. Green, N. I. Vokes, A. M. Feist, M. G. Vander Heiden, C. M. Metallo, *Mol. Cell* **2014**, *55*, 253–263.
[6] J. Ovádi, V. Saks, *Mol. Cell. Biochem.* **2004**, *256*, 5–12.
[7] E. D. Lewandowski, *Biochemistry (Mosc.)* **1992**, *31*, 8916–23.
[8] Z. P. Gerdtzen, in *Genomics Syst. Biol. Mamm. Cell Cult.*, Springer Berlin Heidelberg, Berlin, Heidelberg, **2011**, pp. 71–108.
[9] N. Zamboni, A. Saghatelian, G. J. Patti, *Mol. Cell* **2015**, *58*, 699–706.
[10] S. Niedenführ, W. Wiechert, K. Nöh, *Curr. Opin. Biotechnol.* **2014**, *34C*, 82–90.
[11] C. M. Metallo, J. L. Walther, G. Stephanopoulos, *J. Biotechnol.* **2009**, *144*, 167–174.
[12] J. L. Walther, C. M. Metallo, J. Zhang, G. Stephanopoulos, *Metab. Eng.* **2012**, *14*, 162–71.
[13] K. Schmidt, M. Carlsen, J. Nielsen, J. Villadsen, *Biotechnol. Bioeng.* **1997**, *55*, 831–840.
[14] W. Wiechert, M. Möllney, N. Isermann, M. Wurzel, A. a. De Graaf, *Biotechnol. Bioeng.* **1999**, *66*, 69–85.
[15] J. Choi, M. T. Grossbach, M. R. Antoniewicz, *Anal. Chem.* **2012**, *84*, 4628–32.
[16] C. R. Malloy, a. D. Sherry, F. M. H. Jeffrey, *J. Biol. Chem.* **1988**, *263*, 6964–6971.
[17] A. D. Sherry, C. R. Malloy, P. Zhao, J. R. Thompson, *Biochemistry (Mosc.)* **1992**, *31*, 4833–4837.
[18] T. Cheng, J. Sudderth, C. Yang, A. R. Mullen, E. S. Jin, J. M. Mates, R. J. DeBerardinis, *Proc. Natl. Acad. Sci.* **2011**, *108*, 8674–8679.
[19] M. J. Arauzo-Bravo, K. Shimizu, *Genome Inform. Int. Conf. Genome Inform.* **2001**, *12*, 63–72.
[20] B. Christensen, J. Nielsen, *Metab. Eng.* **1999**, *1*, 282–290.
[21] J. C. Chatham, B. Bouchard, C. Des Rosiers, *Mol. Cell. Biochem.* **2003**, *249*, 105–112.
[22] W. Li, F. Bian, P. Chaudhuri, X. Mao, H. Brunengraber, X. Yu, *NMR Biomed.* **2011**, *24*, 176–187.
[23] S. a. Smith, T. O. Levante, B. H. Meier, R. R. Ernst, *J. Magn. Reson. A* **1994**, *106*, 75–105.
[24] U. Sauer, *Mol. Syst. Biol.* **2006**, *2*, DOI 10.1038/msb4100109.
[25] W. Wiechert, *Metab. Eng.* **2001**, *3*, 195–206.
[26] L. G. Boros, N. J. Serkova, M. S. Cascante, W.-N. P. Lee, *Drug Discov. Today Ther. Strateg.* **2004**, *1*, 435–443.
[27] A. Micchelli, A. Tomassini, C. Puccetti, M. Valerio, G. Peluso, F. Tuccillo, M. Calvani, C. Manetti, F. Conti, *Biochimie* **2006**, *88*, 437–448.
[28] C. Lussey-Lepoutre, K. E. R. Hollinshead, C. Ludwig, M. Menara, A. Morin, L.-J. Castro-Vega, S. J. Parker, M. Janin, C. Martinelli, C. Ottolenghi, et al., *Nat. Commun.* **2015**, *6*, 8784.
[29] J. B. Carrigan, M. A. C. Reed, C. Ludwig, F. L. Khanim, C. M. Bunce, U. L. Günther, *ChemPlusChem* **2016**, *81*, 453–459.
[30] M. A. C. Reed, C. Ludwig, C. M. Bunce, F. L. Khanim, U. L. Günther, *Cancer Metab.* **2016**, *4*, 15.
[31] C. Ludwig, U. L. Günther, *BMC Bioinformatics* **2011**, *12*, 366.
[32] null Gunther, null Ludwig, null Ruterjans, *J. Magn. Reson. San Diego Calif 1997* **2000**, *145*, 201–208.
[33] K. Kazimierczuk, V. Y. Orekhov, *Angew. Chem. Int. Ed Engl.* **2011**, *50*, 5556–5559.
[34] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277–93.

Entry for the Table of Contents (Please choose one layout)

Layout 1:

COMMUNICATION

Text for Table of Contents

((Insert TOC Graphic here))

*Author(s), Corresponding Author(s)**

Page No. – Page No.

Title

Layout 2:

COMMUNICATION

((Insert TOC Graphic here))

*Author(s), Corresponding Author(s)**

Page No. – Page No.

Title

Text for Table of Contents