



# Advanced isotopic labeling for the NMR investigation of challenging proteins and nucleic acids

Jerome Boisbouvier<sup>1</sup> · Lewis E. Kay<sup>2,3</sup>

Published online: 24 July 2018  
© Springer Nature B.V. 2018

With the availability of ultra-high field NMR spectrometers operating at a <sup>1</sup>H frequency of 1 GHz together with the development of new pulse sequences that significantly improve spectral sensitivity and resolution, structural biologists can begin to focus on the study of biological systems that have long been considered beyond the scope of traditional biomolecular NMR spectroscopy. However, without optimized sample preparation strategies it is not possible to exploit all of the possibilities that have become available with the new equipment and the exciting methodologies. The use of advanced isotopic labeling methods, enabling the near 100% incorporation of <sup>1</sup>H, <sup>2</sup>H, <sup>15</sup>N and <sup>13</sup>C spins in strategic positions of target samples with limited scrambling, leading to simplifications of NMR spectra, is required to investigate interactions, structures and dynamics of large and complex biological systems. These sophisticated labeling schemes often allow a significant enhancement in the signal-to-noise of NMR experiments and, concomitantly, a speed-up of their acquisition. A first topical issue of the Journal of Biomolecular NMR on isotope labeling was published 8 years ago (Wagner 2010). The manuscripts assembled in these two issues present the latest developments in the production of NMR samples with optimized labeling schemes to investigate the ever-increasing complex targets of modern structural biology research programs.

Knowledge of the metabolic pathways of the organism used to overexpress target proteins often underlies the

implementation of optimized isotopic labeling strategies. The cover figure of this issue shows a simplified map of the main synthesis and degradation pathways of amino acids in *E. coli*. This map (inspired from the metromap of Lyon) was exploited by Bockmann and collaborators (Lacabanne et al. 2018) to specifically unlabel amino-acids in a target protein sample. In this overview, the authors show that the addition of specific sets of unlabeled amino-acids in standard uniformly (<sup>15</sup>N, <sup>13</sup>C)-labeled minimal media leads to a significant simplification in typically overcrowded 2D <sup>13</sup>C–<sup>13</sup>C solid state NMR spectra of large proteins. The use of irreversible metabolic pathways enables the selective labeling of specific atoms of amino-acids through the addition of a previously synthesized isotopomer of a suitable precursor. In this issue, two papers describe how the metabolic-pathways connecting aromatic residues can be exploited to produce samples with specific labeling of these amino acids. Lichteneker and collaborators (Schörghuber et al. 2018) present a comparison of aromatic precursors entering different stages of the His-, Phe-, Tyr- and Trp-pathways. In a Perspective article, Kainosho and co-workers describe how small quantities of Stereo-Array Isotope Labeled (SAIL, Kainosho et al. 2006) Phe and Trp can be efficiently incorporated into large proteins using *E. coli*, allowing acquisition of 1D and 2D NMR spectra of aromatic residues with an unprecedented resolution (Kainosho et al. 2018).

A second pair of articles describes how atom specific isotope-labeled nucleotides, obtained by chemical and enzymatic synthesis, can be used to improve the quality of NMR spectra of RNAs. Indeed, compared to proteins, NMR studies of nucleic acids are challenged by the presence of only four different nucleotide types, leading to severe overlap in spectra. Carlomagno and collaborators present an overview of segmental-labeling approaches, as well as nucleotide-type and atom-specific labeling of RNAs to simplify solid-state NMR spectra of nucleic acids (Marchanka et al. 2018). The combination of such advanced labeling approaches will contribute to structural investigations of very large RNA

✉ Jerome Boisbouvier  
jerome.boisbouvier@ibs.fr

✉ Lewis E. Kay  
kay@pound.med.utoronto.ca

<sup>1</sup> University of Grenoble Alpes, CNRS, CEA, IBS,  
38000 Grenoble, France

<sup>2</sup> Departments of Molecular Genetics, Biochemistry,  
and Chemistry, University of Toronto, Toronto,  
ON M5S 1A8, Canada

<sup>3</sup> Program in Molecular Medicine, Hospital for Sick Children,  
555 University Avenue, Toronto, ON M5G 1X8, Canada

complexes in the coming years. Complex networks of scalar couplings in sugars and nucleobases also hamper the precise quantification of NMR observables such as relaxation parameters. In the second article Dayie and co-workers introduce ribonucleotides selectively labeled with isolated  $^{13}\text{C}$ – $^1\text{H}$  pairs, enabling the acquisition of accurate relaxation dispersion experiments to characterize conformational exchange in two different RNAs (LeBlanc et al. 2018).

Heterologous expression and labeling of functional and properly folded human protein targets is often not feasible using simple prokaryotic expression systems. Complex machineries are often required to assist protein folding, to incorporate post-translational modifications or to insert the protein of interest into the proper environment, such as a membrane. These challenging proteins, including GPCRs and antibodies, are usually expressed in yeast, insect or mammalian cells. Using algae or yeast extracts as starting labeled material, Grzesiek and co-workers describe a cost-effective and efficient method to produce U- $(^{15}\text{N}, ^{13}\text{C}, ^2\text{H})$  labeled protein in insect cells, enabling the production and the NMR investigation of the G protein-coupled  $\beta$ -adrenergic receptor (Franke et al. 2018). Specific incorporation of  $^{13}\text{CH}_3$  probes in a perdeuterated protein background offers a unique opportunity to simultaneously enhance signal-to-noise and improve spectral resolution of large proteins (Ruschak and Kay 2010; Kerfah et al. 2015). In this issue, Shimada and co-workers propose optimized combinations of perdeuterated and methyl specifically labeled amino-acids to produce proteins in insect cells with labeling of methyl groups. Nevertheless, incorporation of alanine methyl probes at a high level in vivo remains hampered by the presence of transaminases converting endogenous pyruvate into alanine. This limitation was solved by the use of beta-chloro-L-alanine to specifically inhibit alanine transaminase, thus enabling production of the  $\beta$ -adrenergic receptor labeled with  $^{13}\text{CH}_3$  probes distributed along the protein backbone (Kofuku et al. 2018). Kato and co-workers describe an elegant protocol to incorporate methyl specifically labeled leucine in a 150 kDa IgG produced in mammalian cells (Yanaka et al. 2018). Although specifically labeled methyl containing amino acids can be incorporated into proteins using several expression systems, the application of specific labeling in eukaryotic systems remains limited by the cost and availability of specifically labeled amino acids. Yeast expression systems offer an attractive alternative to express and label complex proteins such as GPCRs, as these eukaryotic cells can be grown in 100% deuterated medium using methanol, glycerol or glucose as  $^{13}\text{C}$ - carbon sources. As described by Gardner and co-workers, production and efficient methyl specific labeling of Isoleucine- $\delta 1$  methyl groups in GPCRs can be achieved in yeast using simple 2-ketobutyric precursors (Clark et al. 2018), initially developed for specific labeling of protein overexpressed in *E. coli* (Gardner and Kay

1997). However, the extension of such a simple approach to the labeling of leucine and valine methyl probes using the 2-keto-isovaleric acid precursor remains limited due to its low incorporation level into the target protein. Takahashi and coworkers have performed a systematic investigation of precursor-uptake by *Pichia pastoris*, and have established optimized culture conditions at low pH to increase the incorporation level of the precursor in Valine and Leucine residues (Suzuki et al. 2018). In this manner the number of  $^{13}\text{CH}_3$  probes that are available to investigate important target proteins produced in yeast increases.

Segmental labeling is an alternative approach to reduce overlap in NMR spectra of large proteins. Several established methods are available for segmental isotopic labeling such as intein-mediated ligation, but each has specific requirements and limitations. In this issue, Iwai and co-workers describe an alternative enzymatic approach, using bacterially produced asparagine endopeptidase, and apply it to simplify spectra of armadillo repeat protein, a class of protein particularly difficult to study by NMR due to the large number of repetitive sequences (Mikula et al. 2018). Using intein mediated ligation Meier and collaborators show an elegant application of segmental labeling to the study of a 12  $\times$  59 kDa motor protein (Wiegand et al. 2018). The labeling of the N-terminal segment, in the context of the full-length protein, has allowed the identification by solid-state NMR of the structural elements affected by oligomerization of this *c.a.* 700 kDa protein assembly.

One drawback, commonly observed when producing perdeuterated proteins is the decrease in protein yield, which can affect the project feasibility. Gossert and co-workers present in this special issue optimized protocols for the large scale production of U- $[^{15}\text{N}, ^{13}\text{C}, ^2\text{H}]$  and methyl-specifically labeled proteins using high cell-density fermentation (Klopp et al. 2018). These robust protocols have been extensively tested in a pharmaceutical company for the large scale and parallel production of challenging target proteins, but the key procedures are easily transferred to shake flask cultures for academic applications. To circumvent toxicity of  $\text{D}_2\text{O}$ , Wand and collaborators describe an alternative approach, using  $\text{H}_2\text{O}$ -based M9 medium enriched with a mixture of perdeuterated amino acids and  $^{13}\text{CH}_3$ -precursors to produce large amounts of proteins specifically labeled on Ile, Val and Leu methyls groups (O'Brien et al. 2018).

An additional powerful tool to obtain information on the structure, dynamics and interactions between biological macromolecules involves chemical modifications of the component players. Ubbink and co-workers describe an approach, using caged lanthanoid NMR probes and induced pseudocontact shifts, to detect the displacement of methyl groups upon ligand binding to HSP90 (Lescanne et al. 2018). In the final article of this issue Otting and co-workers describe an approach exploiting amber stop codons

to incorporate phenylalanine substituted in the *para* position with tert-butyl or trimethylsilyl groups during in vitro protein expression (Loh et al. 2018). Due to the fast rotation of the three methyl groups and the proton multiplicity, introduction of these modified amino-acids offers an opportunity to monitor ligand binding in large protein systems using simple 1D NMR experiments.

This special issue has benefited from the valuable input of many participants of the conference and practical school on *Advanced Isotopic Labeling Methods* (AILM) organized every 2 years in Grenoble. Most of the authors and reviewers of these sixteen articles were speakers or instructors. Notably, three trainees of the first AILM practical school are first authors of contributions assembled in this special issue. We would like to thank all of the authors for their contributions and also the reviewers who have offered expertise and advice to improve the quality of the papers in this issue, describing the latest developments in this rapidly evolving field.

## References

- Clark L, Dikiy I, Rosenbaum DM, Gardner KH (2018) On the use of *Pichia pastoris* for isotopic labeling of human GPCRs for NMR studies. J Biomol NMR (in press)
- Franke B, Opitz C, Isogai S, Grahl A, Delgado L, Gossert AD, Grzesiek S (2018) Production of isotope-labeled proteins in insect cells for NMR. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0172-7>
- Gardner KH, Kay LE (1997) Production and incorporation of  $^{15}\text{N}$ ,  $^{13}\text{C}$ ,  $^2\text{H}$  (1H-d1 methyl) isoleucine into proteins for multidimensional NMR studies. J Am Chem Soc 119:7599–7600
- Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Ono AM, Güntert P (2006) Optimal isotope labeling for NMR protein structure determinations. Nature 440:52–57
- Kainosho M, Miyanoiri Y, Terauchi T, Takeda M (2018) Perspective: next generation isotope-aided methods for protein NMR spectroscopy. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0198-x>
- Kerfah R, Plevin MJ, Sounier R, Gans P, Boisbouvier J (2015) Methyl specific isotopic labeling: a molecular tool box for NMR studies of large proteins. Curr Opin Struct Biol 32:113–122
- Klopp J, Winterhalter A, Gébleux R, Scherer-Becker D, Ostermeier C, Gossert AD (2018) Cost-effective large-scale expression of proteins for NMR studies. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0179-0>
- Kofuku Y, Yokomizo T, Imai S, Shiraishi Y, Natsume M, Itoh H, Inoue M, Nakata K, Igarashi S, Yamaguchi H, Mizukoshi T, Suzuki EI, Ueda T, Shimada I (2018) Deuteration and selective labeling of alanine methyl groups of  $\beta_2$ -adrenergic receptor expressed in a baculovirus-insect cell expression system. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0174-5>
- Lacabanne D, Meier BH, Böckmann A (2018) Selective labeling and unlabeled strategies in protein solid-state NMR spectroscopy. J Biomol NMR. <https://doi.org/10.1007/s10858-017-0156-z>
- LeBlanc RM, Longhini AP, Tugarinov V, Dayie TK (2018) NMR probing of invisible excited states using selectively labeled RNAs. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0184-3>
- Lescanne M, Ahuja P, Blok A, Timmer M, Akerud T, Ubbink M (2018) Methyl group reorientation under ligand binding probed by pseudocontact shifts. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0190-5>
- Loh CT, Adams LA, Graham B, Otting G (2018) Genetically encoded amino acids with tert-butyl and trimethylsilyl groups for site-selective studies of proteins by NMR spectroscopy. J Biomol NMR. <https://doi.org/10.1007/s10858-017-0157-y>
- Marchanka A, Kreutz C, Carlomagno T (2018) Isotope labeling for studying RNA by solid-state NMR spectroscopy. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0180-7>
- Mikula KM, Krumwiede L, Plückthun A, Iwai H (2018) Segmental isotopic labeling by asparaginyl endopeptidase-mediated protein ligation. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0175-4>
- O'Brien ES, Lin DW, Fuglestad B, Stetz MA, Gosse T, Tommos C, Wand AJ (2018) Improving yields of deuterated, methyl labeled protein by growing in  $\text{H}_2\text{O}$ . J Biomol NMR. <https://doi.org/10.1007/s10858-018-0200-7>
- Ruschak AM, Kay LE (2010) Methyl groups as probes of supramolecular structure, dynamics and function. J Biomol NMR 46:75–87
- Schörghuber J, Geist L, Platzer G, Feichtinger M, Bisaccia M, Scheibelberger L, Weber F, Konrat R, Lichtenecker RJ (2018) Late metabolic precursors for selective aromatic residue labeling. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0188-z>
- Suzuki R, Sakakura M, Mori M, Fujii M, Akashi S, Takahashi H (2018) Methyl-selective isotope labeling using  $\alpha$ -ketoisovalerate for the yeast *Pichia pastoris* recombinant protein expression system. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0192-3>
- Wagner G (2010) A topical issue: production and labeling of biological macromolecules for NMR investigations. J Biomol NMR 46:1–2
- Wiegand T, Cadalbert R, von Schroetter C, Allain FH, Meier BH (2018) Segmental isotope labelling and solid-state NMR of a  $12 \times 59$  kDa motor protein: identification of structural variability. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0196-z>
- Yanaka S, Yagi H, Yogo R, Yagi-Utsumi M, Kato K (2018) Stable isotope labeling approaches for NMR characterization of glycoproteins using eukaryotic expression systems. J Biomol NMR. <https://doi.org/10.1007/s10858-018-0169-2>