

ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Sensitive proton-detected solid-state NMR spectroscopy of large proteins with selective CH₃ labelling: application to the 50S ribosome subunit

Vilius Kurauskas,^{a,b,c} Elodie Crublet^{*,a,b,c,e}, Pavel Macek,^{a,b,c} Rime Kerfah,^d Diego F. Gauto,^{a,b,c} Jérôme Boisbouvier^{a,b,c} and Paul Schanda^{*,a,b,c}

Solid-state NMR spectroscopy allows the characterization of structure, interactions and dynamics of insoluble and/or very large proteins. Sensitivity and resolution are often major challenges for obtaining atomic-resolution information, in particular for very large protein complexes. Here we show that the use of deuterated, specifically CH₃-labelled proteins result in significant sensitivity gains compared to previously employed CHD₂ labelling, while line widths only marginally increase. We apply this labelling strategy to a 468 kDa-large dodecameric aminopeptidase, TET2, and the 1.6 MDa-large 50S ribosome subunit of *Thermus thermophilus*.

Biomolecular magic-angle spinning (MAS) solid-state NMR (ssNMR) has made substantial progress over the last decade in terms of sensitivity, resolution and information content, and established as an important technique in structural biology, addressing structure and dynamics in systems with increasing complexity.¹⁻³ The progress was enabled by improved NMR instrumentation, in particular probes allowing fast MAS (up to 100 kHz), improved pulse sequences, as well as better sample preparation and isotope labelling techniques. Experiments based on proton-detection are particularly powerful for the study of protein structure, interactions and dynamics. The large gyromagnetic ratio of protons translates to high intrinsic detection sensitivity, and ability to measure long-range distances. However, the large ¹H gyromagnetic ratio and the corresponding strong ¹H-¹H dipolar couplings also result in large ¹H line widths. Proton-detected experiments are, therefore, particularly powerful when protons are introduced

only at selected sites of interest in an otherwise deuterated environment, and when high MAS frequencies are employed, leading to efficient averaging of ¹H-¹H interactions. Methyl groups are particularly interesting reporters of protein structure and dynamics in proteins, due to their widespread abundance in proteins, often located in important hydrophobic core regions⁴. In solution-state NMR, methyl-directed studies have provided a wealth of information about dynamics, interactions and structure in particular for large complexes (up to 1 MDa), because in this case methyls are the only observable NMR reporters.^{5,6} Protocols for introducing methyls of all iso-topomers (CH₃, CH₂D, CHD₂) into otherwise deuterated proteins⁴ have been established. The three isotopomers differ in the number and nature of interactions, and thus in the spin relaxation properties, sensitivity and resolution. In solution-state NMR, the most commonly employed isotopomer is CH₃, which generally provides the highest sensitivity due to the three-fold spin multiplicity. As long as optimized pulse schemes are used⁷, the spectra of all isotopomers exhibit similar line widths.⁸ In ssNMR, only few studies have reported the use of specifically methyl-labelled samples for studying dynamics and structure.^{3,6-8} All these studies have employed CHD₂ labelling, with the exception of a brief note of CH₃ labelling.⁹ The primary argument for using the CHD₂ isotopomer is related to the fact that there are no intra-methyl ¹H-¹H dipolar couplings in CHD₂, and that these interactions are expected to lead to strong line broadening; this assumption was supported by pioneering experiments on stochastically labelled proteins at low (20 kHz) MAS frequency,¹⁰ but not systematically investigated. Here we show that, in experiments performed at high MAS frequency (50-60 kHz) the line widths in CHD₂ and CH₃ methyl are comparable, and that CH₃ labelling results in improved sensitivity compared to CHD₂. The enhanced sensitivity opens possibilities to study even very large objects, as demonstrated with the 1.6 MDa-large 50S ribosome, where sensitivity is a great challenge.

- a. Université Grenoble Alpes, Institut de Biologie Structurale, Grenoble.
b. CEA, Institut de Biologie Structurale, F-38044 Grenoble, France.
c. CNRS, Institut de Biologie Structurale, F-38044 Grenoble France.
d. NMR-Bio. 5 place Robert Schumann, F-38025 Grenoble, France.
e. Present address: NMR-Bio. 5 place Robert Schumann, F-38044 Grenoble, France.

e-mail: crublet@nmr-bio.com and schanda@ibs.fr

Electronic Supplementary Information (ESI) available: comprehensive experimental details. See DOI: 10.1039/x0xx00000x

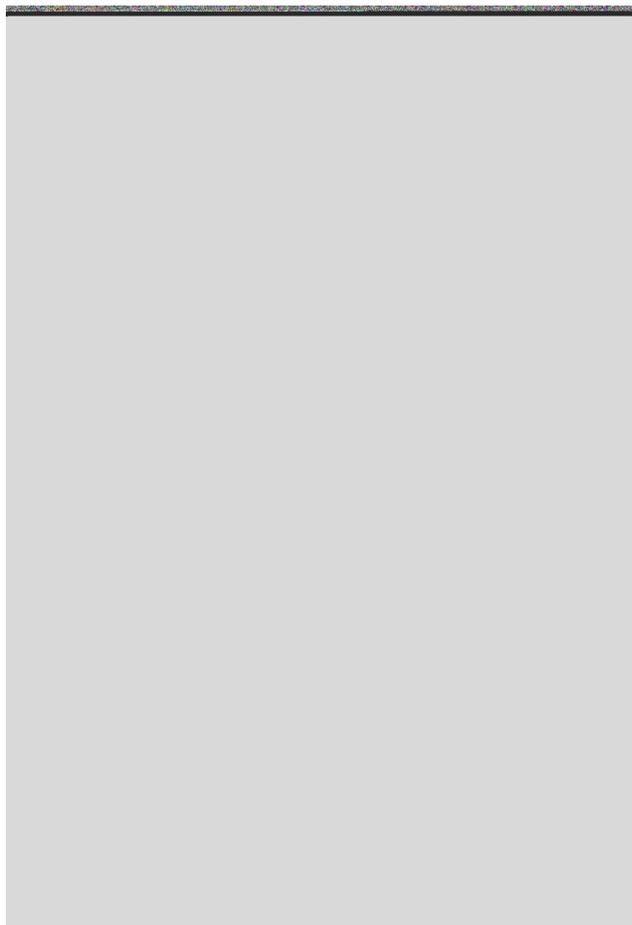


Fig. 1. Proton-detected 2D methyl ^1H - ^{13}C spectra of CH_3 (blue) and CHD_2 (red) methyls of (a) isoleucines and (b) alanines in TET2, obtained at 14.1 T (600 MHz ^1H Larmor frequency). These spectra were obtained with the CP-based pulse sequence shown in ESI Figure S1a. Traces along the positions indicated with dashed lines and numbers are shown in (c)-(e). Dashed 1D traces were obtained at the indicated positions in the spectra, but different pulse sequences, either omitting ^2H decoupling, or using MQ evolution (see Figure S1). Line-widths (full-width at half maximum) are indicated for these selected peaks. (f) shows the ratio of line widths observed for CH_3 and CHD_2 correlations, merged for both Ala and Ile samples. (g) Intensity ratios for cross-peaks observed in the Ile spectrum, taking into account the protein amounts in CH_3/CHD_2 samples. For all these spectra, resolution enhancement was applied using a 70-degree shifted sine-bell square function (90 degrees would correspond to a cosine, i.e. no resolution enhancement), see ESI.

In order to investigate the properties of the different methyl isotopomers in ssNMR, we compared four samples of the enzyme TET2 from *Pyrococcus horikoshii*, a dodecameric (12x39 kDa) aminopeptidase involved in the cellular protein quality control, with different methyl labelling schemes. (Details about sample preparation can be found in the ESI.) With 353 residues per monomer and a total oligomer mass of 468 kDa, TET2 is among the largest proteins studied so far by ssNMR, making it thus a considerable challenge in terms of resolution and sensitivity. In the samples we incorporated either CH_3 or CHD_2 methyl groups on isoleucine residues or on alanines, respectively, in an otherwise deuterated background. The different lengths and flexibility of Ala and Ile side chains

generally lead to different dynamics, reflecting the range of methyl cases encountered in proteins. Figure 1a,b shows comparisons of ^1H - ^{13}C correlation spectra of these four samples, obtained with cross-polarization transfer steps at a MAS frequency of 57 kHz. High-resolution spectra are obtained for this very large protein assembly for both isotopomers and both residue types. The spectra are similar to solution-state spectra of TET, showing that structure and dynamics in our solid-state preparations are conserved in the solid state, and allowing us to transfer previously known assignments from solution-state NMR¹¹ (see Supporting Information Figure S2). In CHD_2 groups, the observed line widths are of the order of 30 – 45 Hz (^1H) and 20-30 Hz (^{13}C), respectively (Figure 1c-e). These values are in good agreement with previously reported line widths obtained in fast-MAS solid-state NMR experiments on smaller model proteins with CHD_2 labelling^{9,12,13}. For the CH_3 isotopomer, the line widths along the ^{13}C dimension are similar to the ones observed with CHD_2 groups (20-30 Hz), suggesting that the intra-methyl heteronuclear ^1H - ^{13}C dipolar coupling is not the dominant factor for line width under the chosen conditions (heteronuclear decoupling, fast MAS). This is expected, given that fast-MAS, the high deuteration level of the surroundings, and the 3-fold reduced dipolar ^1H - ^{13}C coupling that results from the rapid methyl rotation, make decoupling very efficient. Remarkably, the ^1H line widths in the CH_3 sample are only slightly larger than the ones in CHD_2 samples (Figure 1e,f). The fact that reducing the number of methyl proton spins from three (CH_3) to one (CHD_2) has only a modest effect on line widths suggests that the intra-methyl dipolar couplings are not determining the ^1H line widths. Sample inhomogeneity, dipolar inter-actions to remote protons as well as dynamics may be important contributions to line widths, such that intra-methyl ^1H - ^1H dipolar interactions do not dominate. Of note, given that the line widths observed here are similar to microcrystalline samples,^{9,12} our conclusions hold also for those systems. In more heterogeneous systems (fibrils, membrane proteins) the differences in line widths between the two isotopomers are expected to be similar or smaller than here, because line widths due to sample heterogeneity is independent of the labelling scheme. Importantly, in terms of sensitivity the CH_3 labelling scheme outcompetes CHD_2 labelling, on average by a factor of about 1.9 (Figure 1g). This is due to the proton spin multiplicity (expected gain of 3), reduced by the somewhat faster relaxation of CH_3 methyls (see below). The experiments shown in Figure 1a,b were recorded with ^2H decoupling during chemical-shift evolution periods, which requires a probehead able to tune to ^2H frequency. Such equipment is not standard in contemporary MAS probes. As expected, while omission of ^2H decoupling has virtually no effect on line widths and sensitivity in CH_3 -labelled methyls (Figure 1c), due to the absence of $^1\text{J}_{\text{CD}}$ couplings, in CHD_2 groups the undecoupled ^{13}C line widths are dramatically increased due to the scalar coupling (~ 20

Hz) multiplet structure.¹² Therefore, if ²H decoupling is impossible with the available hardware, the CH₃ isotopomer is clearly favoured in terms of resolution and sensitivity. The effects of ²H decoupling on ¹H line widths are negligible, as the ²J_{HD} coupling is small (~1-3 Hz).¹² We have also investigated two different types of ¹³C evolution, single-quantum (SQ), and multiple-quantum (MQ) evolution. In solution-state NMR, in the case of slow overall molecular tumbling, MQ coherence life times are significantly longer than SQ life times¹⁴. In contrast, Figure 1d establishes that MQ lines in our ssNMR experiments are almost two-fold larger than their SQ counterparts. This can be understood by considering that due to the absence of overall tumbling, cross-correlated relaxation is small (reducing the potential benefit of MQ evolution), and dephasing due to remote protons outweighs the potential gains from cross-correlated relaxation. We also investigated the relative sensitivity obtained with INEPT-transfer and CP-transfer. INEPT-based transfer is favoured by large-scale dynamics and extensive deuteration, because both factors increase coherence life times and thus INEPT transfer efficiency. We find indeed that for the more dynamic Ile residues INEPT outcompetes CP-transfer, while for Ala we find similar sensitivity (Figure S4).

Line widths are related to the life times of spin coherences, which are determined by spin relaxation (i.e., dynamics) and dephasing that is due to incompletely averaged dipolar interactions; the latter contribution depends on the MAS frequency. To further gain insight into the properties of the different labellings, we performed systematic analyses of apparent relaxation rate constants. These measurements, shown in Figure 2, corroborate the trends shown in Figure 1: (i) ¹H R₂' relaxation rates are similar in CH₃ and CHD₂ samples, with values of ~100 s⁻¹ and ~80 s⁻¹, respectively, at high MAS frequency (55-57 kHz). The expected line widths from these numbers (R₂'/π, ~25-30 Hz) are smaller than the experimentally observed values (~35-45 Hz), showing that part of the line width originates from sample and/or magnetic-field heterogeneity. (ii) ¹³C R₂' of ~20-40 s⁻¹ are found at high MAS frequencies (~55 kHz). The line widths predicted from these numbers (6-12 Hz) are significantly smaller than the experimentally observed line widths in this study and crystalline ubiquitin before¹² (~18-25 Hz), showing that heterogeneity limits the resolution of ¹³C lines. (iii) Omission of ²H decoupling leads to significantly shorter ¹³C T₂' in CHD₂ (but not CH₃), in agreement with the larger line widths (Figure 1c) (iv) MQ coherences decay faster than their SQ counterpart, again in line with data shown in Figure 1d. (v) Furthermore, proton R₁ rate constants are ~15% higher in CH₃ compared to CHD₂. This allows higher repetition rate of experiments, and thus slightly increased sensitivity for CH₃ samples (see Figure S4). (vi) We also measured ¹³C R₁ and ¹³C R_{1ρ} rate constants, and find that they are not MAS-frequency dependent. This finding establishes that coherent dephasing (i.e. contribution to the decay that is not due to dynamics, and

which would be MAS dependent) is negligible, allowing quantitative interpretation of methyl dynamics. While ¹³C R₁ has been used for this purpose before,¹³ the possibility to quantitatively interpret R_{1ρ} enables the study of slow (ns-ms) motion.¹⁵

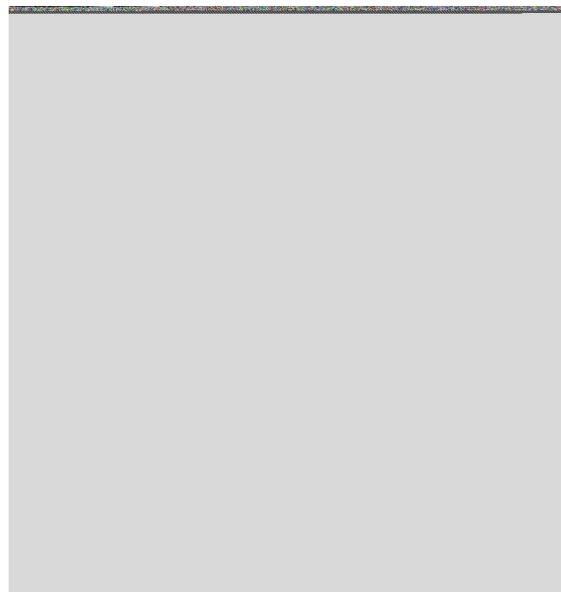


Fig. 2. MAS-frequency dependence of various spin relaxation parameters in Ile-labelled (circles) and Ala-labelled (triangles) TET2, with CH₃ (blue) and CHD₂ (red) isotopomers. Bulk methyl relaxation rate constants were obtained from a series of 1D spectra, using pulse sequences shown in ESI Figure S1.

Having thus established CH₃ labelling provides enhanced sensitivity compared to CHD₂, with similar resolution, we turned to a very challenging application, the 50S ribosome particle from *Thermus thermophilus*. Ribosomes are among the most complex biological machineries, composed of RNAs and proteins, with a total molecular mass of 1.6 MDa (for the 50S subunit). Unlike TET2, where the 12-fold symmetry improves sensitivity and reduces spectral complexity, ribosomes are not symmetric. Sensitivity and resolution are thus a great challenge. Earlier studies by solution-state NMR have observed only highly mobile isotope-labelled proteins attached to ribosomes¹⁶, or flexible nascent polypeptides exiting the ribosome,^{17,18} while less flexible proteins are invisible due to unfavourable relaxation properties. Recently, ssNMR spectra have been reported for tightly ribosome-bound proteins, which were isotope labelled and bound to unlabelled 50S and 30S subunits, respectively.^{19,20} This selective labelling of a single protein overcomes the severe resonance overlap problem but requires reconstitution of ribosomes from labelled/unlabelled subunits, which is not generally possible. We investigate here the possibility to study uniformly labelled ribosome (50S subunit, 1.6 MDa), obtained from bacterial cells grown in uniformly deuter-

COMMUNICATION

Journal Name

ated, Ala-CH₃-labeled media, and sedimentation into a 1.3 mm ssNMR rotor²¹. The HSQC-type ¹H-¹³C correlation spectrum obtained in a total experimental time of 14 hours is shown in Figure 3 (black). As expected, the large number of non-equivalent Ala methyl sites – 337 in total – results in significant resonance overlap. Nonetheless, for the resolved cross-peaks the line widths are similar to TET2 (~25-30 Hz for ¹³C and 40-50 Hz for ¹H). Signal-to-noise ratios of these cross-peaks are of the order of 3 to 8. Given the small amount of sample (1.2 nmoles) these results are encouraging, and demonstrate that methyl labelling is a route to gain residue-specific information of very large complexes. We also performed an additional experiment, using a shorter ¹³C evolution time and larger number of scans to favour sensitivity and confirm that the detected cross peaks are indeed reproducible (red spectrum in Figure 3). Resonance assignment and the spectral complexity are of course challenging. Solutions may involve unlabelling of protein components that can be reconstituted, or chemically methyl-labelled amino acid with low abundance, such as methylated cysteines²² or lysines. The data shown here – to our knowledge the first reported ssNMR spectra of uniformly methyl-labelled ribosomes – establish that in terms of sensitivity and line widths, studies of ribosomes are feasible.



Fig. 3. Alanine methyl ssNMR spectra of *Thermus thermophilus* ribosome, obtained with 1.5 mg of deuterated, Ala-¹³CH₃ labelled material. (a) 2D spectra obtained with INEPT-transfer (black, pulse sequence Figure S1E), obtained in a total time of 14 hours. The spectrum in red was obtained in 9 hours, using CP-transfers (sequence Figure S1A) with increased number of scans at the expense of resolution (see ESI for details). (b) 1D traces of the black spectrum along the ¹H and ¹³C dimensions were extracted at the positions indicated with dashed lines, and the colours of the traces correspond to the colours of the arrows.

In conclusion, we have shown here that CH₃-methyl-labelling allows for high-resolution proton-detected ssNMR at high MAS frequencies. In terms of sensitivity, CH₃ labelling significantly outcompetes CHD₂ labelling. We believe that CH₃ methyl labelling will be very useful for further ssNMR studies of structure, dynamics and

interactions, in particular for large systems, as exemplified here.

This work was supported by the European Research Council (ERC-StG-2012-311318 and ERC-StG-2010-260887) and used the platforms of the Grenoble Instruct Center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB).

- 1 A. Loquet, N. G. Sgourakis, R. Gupta, K. Giller, D. Riedel, C. Goosmann, C. Griesinger, M. Kolbe, D. Baker, S. Becker and A. Lange, *Nature*, 2012, **486**, 276–9.
- 2 C. Wasmer, A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek and B. H. Meier, *Science*, 2008, **319**, 1523–6.
- 3 L. B. Andreas, M. Reese, M. T. Eddy, V. Gelev, Q. Z. Ni, E. A. Miller, L. Emsley, G. Pintacuda, J. J. Chou and R. G. Griffin, *J. Am. Chem. Soc.*, 2015, **137**, 14877–14886.
- 4 R. Sprangers, A. Velyvis and L. E. Kay, *Nat. Methods*, 2007, **4**, 697–703.
- 5 A. M. Ruschak and L. E. Kay, *J. Biomol. NMR*, 2010, **46**, 75–87.
- 6 D. Sheppard, R. Sprangers and V. Tugarinov, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **56**, 1–45.
- 7 V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw and L. E. Kay, *J. Am. Chem. Soc.*, 2003, **125**, 10420–10428.
- 8 J. E. Ollerenshaw, V. Tugarinov, N. R. Skrynnikov and L. E. Kay, *J. Biomol. NMR*, 2005, **33**, 25–41.
- 9 L. B. Andreas, T. Le Marchand, K. Jaudzems and G. Pintacuda, *J. Magn. Reson.*, 2015, **253**, 36–49.
- 10 V. Agarwal, A. Diehl, N. Skrynnikov and B. Reif, *J. Am. Chem. Soc.*, 2006, **128**, 12620–12621.
- 11 C. Amero, M. Asunción Durá, M. Noirclerc-Savoie, A. Perollier, B. Gallet, M. J. Plevin, T. Vernet, B. Franzetti and J. Boisbouvier, *J. Biomol. NMR*, 2011, **50**, 229–236.
- 12 M. Huber, O. With, P. Schanda, R. Verel, M. Ernst and B. H. Meier, *J. Magn. Reson.*, 2012, **214**, 76–80.
- 13 V. Agarwal, Y. Xue, B. Reif and N. R. Skrynnikov, *J. Am. Chem. Soc.*, 2008, **130**, 16611–16621.
- 14 J. E. Ollerenshaw, V. Tugarinov and L. E. Kay, *Magn. Reson. Chem.*, 2003, **41**, 843–852.
- 15 P. Schanda and M. Ernst, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2016, **96**, 1–46.
- 16 F. A. A. Mulder, L. Bouakaz, A. Lundell, M. Venkataramana, A. Liljas, M. Akke and S. Sanyal, *Biochemistry*, 2004, **43**, 5930–5936.
- 17 L. D. Cabrita, et al, *Nat. Struct. Mol. Biol.*, 2011, **18**, 278–285.
- 18 S.-T. D. Hsu, P. Fucini, L. D. Cabrita, H. Launay, C. M. Dobson and J. Christodoulou, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 16516–16521.
- 19 E. Barbet-Massin, C. T. Huang, V. Daebel, S. T. D. Hsu and B. Reif, *Angew. Chemie - Int. Ed.*, 2015, **54**, 4367–4369.
- 20 I. Gelis, V. Vitzthum, N. Dhimole, M. A. Caporini, A. Schedlbauer, D. Carnevale, S. R. Connell, P. Fucini and G. Bodenhausen, *J. Biomol. NMR*, 2013, **56**, 85–93.
- 21 I. Bertini, C. Luchinat, G. Parigi, E. Ravera, B. Reif and P. Turano, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 10396–10399.
- 22 T. L. Religa, A. M. Ruschak, R. Rosenzweig and L. E. Kay, *J. Am. Chem. Soc.*, 2011, **133**, 9063–9068.