



HAL
open science

An optimized isotopic labelling strategy of isoleucine- γ 2 methyl groups for solution NMR studies of high molecular weight proteins

Isabel Ayala, Olivier Hamelin, Carlos D Amero, Ombeline Pessey, Michael J Plevin, Pierre Gans, Jérôme Boisbouvier

► To cite this version:

Isabel Ayala, Olivier Hamelin, Carlos D Amero, Ombeline Pessey, Michael J Plevin, et al.. An optimized isotopic labelling strategy of isoleucine- γ 2 methyl groups for solution NMR studies of high molecular weight proteins. *Chemical Communications*, 2012, 48 (10), pp.1434-1436. <10.1039/C1CC12932E>. <hal-01132447>

HAL Id: hal-01132447

<https://hal.science/hal-01132447v1>

Submitted on 22 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Cite this: DOI: 10.1039/c1cc12932e

www.rsc.org/chemcomm

COMMUNICATION

An optimized isotopic labelling strategy of isoleucine- γ_2 methyl groups for solution NMR studies of high molecular weight proteins \dagger Isabel Ayala, \S^a Olivier Hamelin, \S^b Carlos Amero, ac Ombeline Pessey, a Michael J. Plevin, a Pierre Gans d and Jérôme Boisbouvier *a

Received 18th May 2011, Accepted 14th July 2011

DOI: 10.1039/c1cc12932e

An efficient synthetic route is proposed to produce 2-hydroxy-2-ethyl-3-oxobutanoate for the specific labelling of Ile methyl- γ_2 groups in proteins. The ^2H , ^{13}C -pattern of the biosynthetic precursor has been designed to optimize magnetization transfer, in large proteins, between these important structural probes and their corresponding backbone nuclei.

The use of solution NMR spectroscopy to analyze the dynamics, interactions and function of large proteins (>100 kDa) is becoming increasingly feasible. Key to this advance has been the development of protocols for the expression of selectively methyl-protonated perdeuterated proteins.¹ Methyl groups have been proven to be ideal molecular probes for solution NMR spectroscopy studies of large proteins. Several protocols for selective protonation of Ile- δ_1 ,² Leu- $\delta_{1/2}$ and Val- $\gamma_{1/2}$,³ Ala- β ,⁴ and Met- ϵ ⁵ methyl groups have been reported. In addition, we have recently outlined a protocol for the stereospecific protonation of Leu- δ and Val- γ methyl groups.⁶ In each of these protocols an isotope-labelled metabolic precursor or amino acid is added to a perdeuterated expression culture, which is then incorporated into the target protein by the bacteria without detectable scrambling. NMR signals of protonated methyl groups can be exploited as sensitive site-specific probes for the characterisation of intermolecular interactions⁷ or the detection of important structural restraints.⁸ Methyl-protonated perdeuterated proteins can also be used to report molecular dynamics on a wide range of time-scales.⁹

One drawback of perdeuteration, even when combined with methyl-selective protonation, is the unavoidable reduction in the number of sites that can be monitored by NMR spectroscopy. It is possible to overcome this limitation by using several complementary methyl-protonation strategies.¹ In this way data can

be collected from several samples with different methyl-labelling patterns which, when considered together, would provide a more comprehensive analysis of structure, function or dynamics. It is therefore highly beneficial to be able to specifically and selectively protonate as many different types of methyl groups as possible. Ile has two methyl groups though, to date, the major focus of isotope-labelling strategies has been the δ_1 -position.² Recently, a protocol for regioselective isotope-labelling of the γ_2 -methyl group of Ile was proposed¹⁰ that uses derivatives of 3-keto-acids.⁶ However, a synthetic scheme for the production of this new bioorganic precursor has not yet been described. Moreover, due to the chemical composition of the 3-keto acid used, the reported scheme produces Ile residues in which only the γ_2 position is [^{13}C]-enriched. Resonance assignment of methyl groups in large proteins is commonly achieved by correlating the methyl group signal with previously assigned signals of backbone nuclei. To efficiently connect these nuclei using triple resonance NMR experiments requires an isotope labelling scheme that produces a linear ^{13}C spin system in the amino acid side chain.¹¹ In this regard, the recently reported labelling protocol is not optimal as the γ_2 -methyl group is isolated from nuclei in the protein backbone.

Here, we report a scheme for the chemical synthesis of an isotopically-labelled metabolic precursor that enables selective biosynthetic incorporation of γ_2 -methyl protonated Ile into perdeuterated proteins. The resulting ^{13}C -labelling pattern has been optimized to allow the protonated methyl group to be efficiently connected to nuclei in the backbone.

The labelling strategy outlined here is based on 2-(*S*)-2-hydroxy-2-ethyl-3-oxobutanoate, a biosynthetic precursor of Ile in *E. coli*, which can be conveniently synthesized as described in Fig. 1 and Fig. S1 (ESI \ddagger). Ethyl-labelled 2-hydroxy-2-ethyl-3-oxobutanoate was prepared in three steps directly from commercially available materials with a yield of 79%. Ethyl-protected [U^{-13}C]-3-oxobutanoate (**I**) was initially treated with potassium carbonate in the presence of deuterated iodoethane. The resulting product (**II**) was then selectively oxidized by dimethyldioxirane in the presence of Ni(II) ions as catalysts to produce ethyl 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate. Deprotection of the product was performed in H_2O at pH 13 to produce specifically-labelled 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate (**III**) which was added to an M9/D $_2\text{O}$ -based *E. coli* culture

^a CNRS, Institut de Biologie Structurale, Grenoble, France.

E-mail: jerome.boisbouvier@ibs.fr

^b Université de Grenoble, Laboratoire de Chimie et Biologie des Métaux, 17 avenue des Martyrs, 38054 Grenoble Cedex 9, France

^c Centro de Investigaciones Químicas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Mexico

^d CEA, IBS-41 rue J. Horowitz, F-38027 Grenoble Cedex 1, France

\dagger This article is part of the ChemComm 'Emerging Investigators 2012' themed issue.

\ddagger Electronic supplementary information (ESI) available: Experiment procedures and NMR spectra. See DOI: 10.1039/c1cc12932e

\S These authors contributed equally to this work.

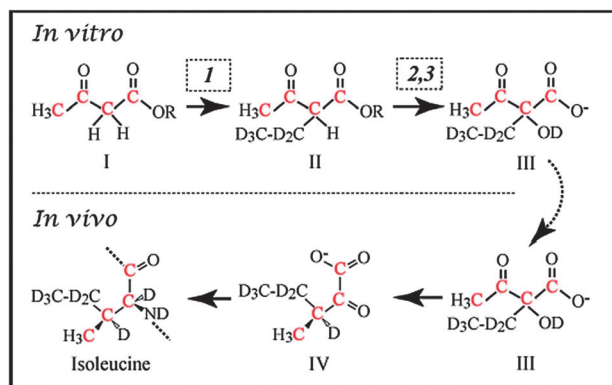


Fig. 1 Simplified protocol for the production of specifically [$^{13}\text{C}_3$]-Ile- γ_2 -labelled proteins. I: [$^{13}\text{C}_3$]-3-oxobutanoate (protected); II: 2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate (protected); III: 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate; IV: (*S*)-3-[^{13}C]methyl-2-oxo-[5- $^2\text{H}_3$,4- $^2\text{H}_2$][1,2,3- $^{13}\text{C}_3$]pentanoate. Step 1: K_2CO_3 , $\text{CD}_3\text{CD}_2\text{I}$; step 2: dimethyldioxirane, $\text{Ni}(\text{OAc})_2$; step 3: NaOH , pH 13 (see text and Fig. S1 (ESI ‡)) for details). ^{13}C atoms are indicated in red. *In vivo*: the precursor is added to a $^2\text{H}/^{12}\text{C}/^{15}\text{N}$ M9 media 1 h prior to induction and is subsequently incorporated into Ile sites in the overexpressed protein (Fig. S1, ESI ‡).

medium without further purification. The specifically-labelled 2-(*S*) stereo-isomer of 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate is converted *in vivo* to (*S*)-3-[^{13}C]methyl-2-oxo-[5- $^2\text{H}_3$,4- $^2\text{H}_2$][1,2,3- $^{13}\text{C}_3$]pentanoate (**IV**) by the *E. coli* enzymes ketolacid reductoisomerase (EC 1.1.1.86) and dihydroxy-acid dehydratase (EC 4.2.1.9), before being incorporated into the Ile side-chain (Fig. 1 and Fig. S1 (ESI ‡)). Saturation of the Ile metabolic pathway with exogenous 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate was achieved by the addition of the specifically-protonated precursor at a concentration higher than 100 mg L^{-1} of perdeuterated culture medium, one hour prior to induction of protein overexpression (Fig. S2, ESI ‡). Surprisingly, low-level isotope scrambling of the precursor was observed at the *proR* methyl positions of Leu and Val. This observation remains difficult to explain considering the current understanding of Ile and Leu/Val metabolic pathways in *E. coli*. Practically, however, scrambling to *proR* sites was efficiently eliminated by adding saturating amounts (200 mg L^{-1}) of perdeuterated isovalerate, a precursor of Leu and Val (Fig. S3, ESI ‡).

The benefits of using methyl-specific protonation can be most clearly seen in NMR spectra of large perdeuterated proteins. The Ile- γ_2 labelling strategy was applied to malate synthase G, a 723 amino acid single chain protein that contains 44 Ile residues.¹² Fig. 2 shows a 2D methyl-TROSY spectrum¹³ of U-[^2H , ^{12}C , ^{15}N], U-[^{13}C , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, $^{13}\text{C}^1\text{H}_3$ - γ_2]Ile-labelled MSG which was expressed from *E. coli* supplemented with 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate. Using this labelling scheme, it was possible to acquire excellent quality, well-resolved NMR spectra of MSG. The spectra show no evidence for CH_2D or CD_2H isotopomers demonstrating that H/D exchange of the γ_2 -methyl group does not occur following the introduction of the precursor to an M9/ D_2O -based expression medium. Furthermore, no signals were observed for Ile- δ_1 methyl groups and isotope enrichment was specific for Ile- γ_2 labelling (*i.e.* no unwanted labelling of other methine, methylene or methyl sites). This specificity means

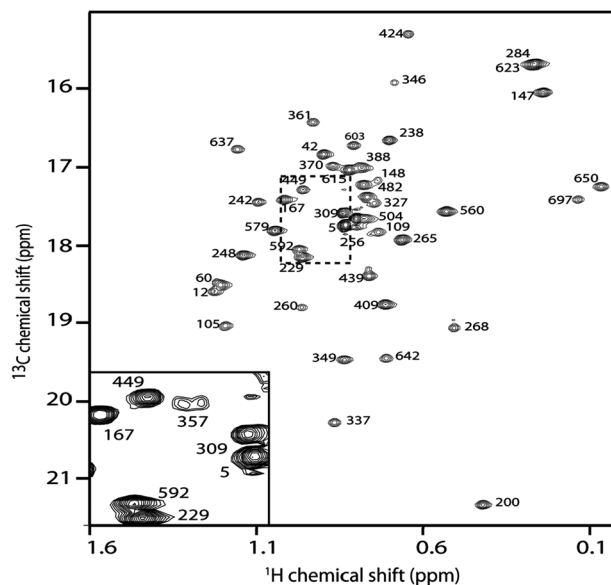


Fig. 2 CT-methyl-TROSY spectra acquired for a 1 mM U-[^{15}N , ^{12}C , ^2H], U-[^{13}C , $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, $^{13}\text{C}^1\text{H}_3$ - γ_2]Ile-labelled MSG sample prepared using 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate. The spectrum was acquired at 37°C on an 800 MHz spectrometer. Each methyl signal has been assigned using “out-and-back” HCC and HC(C)C experiments which connect Ile- γ_2 resonances to the sequentially assigned $\text{C}\alpha$ and $\text{C}\beta$ resonances.¹⁷ The bottom left box displays an expansion of the region of the spectra (delimited by the dashed box) at a lower contour level to show weak methyl- γ_2 cross-peak of Ile-357 (the assignment is based on the analysis of ^{13}C -edited NOESY experiment—Fig. S4 (ESI ‡)).

that 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate can be used in combination with other methyl labelling schemes.^{2–5} Thus, when used in combination with perdeuterated isovalerate, specifically-labelled 2-hydroxy-2-ethyl-3-oxobutanoate is an attractive metabolic precursor for efficient [^1H , ^{13}C]-labelling of Ile- γ_2 methyl groups in fully perdeuterated proteins.

The protocol presented here has several advantages over previously reported schemes.^{14,15} First, only the Ile- γ_2 methyl groups are observed (Fig. 2), which removes the possibility of signal superimposition with Ala- β , Thr- γ_2 , Ile- δ_1 and Val- $\gamma_{1/2}$ resonances in this overcrowded region of 2D methyl-TROSY spectra.¹⁵ Second, as the level of protonation at the Ile- γ_2 methyl position is nearly 100%, a gain in sensitivity of a factor 4 to 6 is expected compared to partially perdeuterated samples.^{14,15} Moreover, more simple and sensitive pulse sequences can be used as it is not necessary to filter out signals from unwanted isotopomers (*i.e.* $^{13}\text{CH}_3$, $^{13}\text{CH}_2\text{D}$ or $^{13}\text{CHD}_2$).¹⁶

The labelling pattern of the precursor described here has been designed to enable the methyl- γ_2 frequencies to be connected to carbon and nitrogen nuclei in the backbone *via* “out-and-back” triple-resonance experiments. The [^{13}C]-labelling pattern produced when using 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate “linearizes” the isoleucine side chain.¹¹ A protein produced from perdeuterated M9 media ($\text{D}_2\text{O} > 99.8\%$, U-[^2H]-glucose, $^{15}\text{ND}_4\text{Cl}$) supplemented with saturating amounts of 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]-3-oxobutanoate will be uniformly perdeuterated with only the Ile γ_2 -positions protonated ($^{13}\text{C}^1\text{H}_3$ isotopomer only). The $\text{C}\gamma_2$, C' , $\text{C}\alpha$ and $\text{C}\beta$ positions of Ile will be [^{13}C]-labelled.

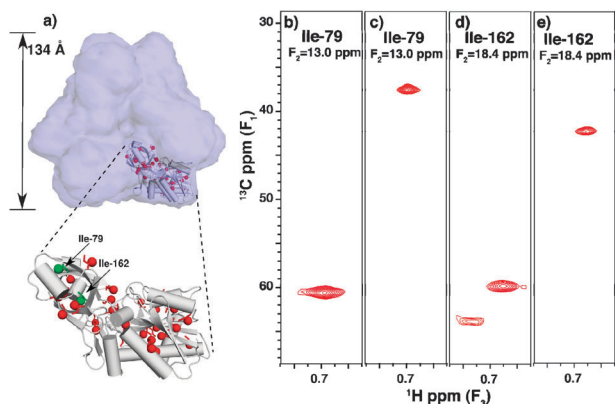


Fig. 3 (a) Quaternary structure of the 468 kDa dodecameric PhTET2 particle. On the bottom left side is presented an expansion of the tertiary structure of the monomeric subunit with the position of the isoleucine- γ_2 methyl probes depicted by red sphere. (b to e) 2D-(F_1, F_3) extracts of 3D HCC (c and e) and HC(C)C experiments (b and d) correlating in F_1 C β (HCC) or C α (HC(C)C) with C γ_2 (F_2) and H γ_2 (F_3) frequencies of PhTET2 Ile-79 (b and c) and Ile-162 (d and e). Both spectra have been recorded at 50 °C in D $_2$ O, using a 1 mM sample U-[$^{15}\text{N}, ^{12}\text{C}, ^2\text{H}$], U-[$^{13}\text{C}'$, $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, $^{13}\text{C}^1\text{H}_3$ - γ_2]Ile-labelled, on NMR spectrometers operating at proton frequencies of 600 MHz (HC(C)C) and 800 MHz (HCC).

Only the C γ_1 and C δ_1 positions of Ile, as all other carbon sites in the protein, will possess ^{12}C atoms.

Sequence specific assignments of Ile γ_2 -methyl resonances of MSG (82 kDa) were obtained using “out-and-back” HCC and HC(C)C-relay experiments recorded for a 1 mM sample of protein labelled using 2-hydroxy-2-[$^2\text{H}_5$]ethyl-[1,2,3,4- $^{13}\text{C}_4$]3-oxobutanoate. Using these two NMR spectra and previously reported C β and C α resonance assignments¹⁷ it was possible to assign 43 of the Ile- γ_2 methyl resonances detected in 2D methyl-TROSY spectra. Only the γ_2 -methyl resonance of Ile-357 was not detected in 3D spectra due to severe line broadening of the methyl resonance (Fig. 2). Nevertheless, it was possible to assign the resonance frequency of this last Ile methyl group to a weak peak in the 2D-methyl TROSY by using NOEs to neighbouring assigned methyl groups (Fig. S4, ESI†).

In order to evaluate the utility of this labelling strategy for very large proteins we have applied it to a multimeric aminopeptidase (PhTET2) involved in polypeptide degradation in the hyperthermophilic Archaea, *Pyrococcus horikoshii*. PhTET2 assembles to form a 468 kDa homododecameric particle with tetrahedral shape.¹⁸ Each of the twelve 39 kDa subunits contains 34 isoleucines. A 1 mM sample of U-[$^2\text{H}, ^{12}\text{C}, ^{15}\text{N}$], U-[$^{13}\text{C}'$, $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, $^{13}\text{C}^1\text{H}_3$ - γ_2]Ile-labelled TET2 was prepared and 2D methyl-TROSY, 3D HCC and 3D HC(C)C NMR experiments were acquired. Examples of NMR spectra are presented in Fig. 3. Despite the large size of PhTET2—characterized by an overall tumbling time of *ca.* 0.2 μs at 50 °C—it was still possible to obtain high quality NMR spectra (Fig. 3). Analysis of triple resonance experiments recorded on an Ile- γ_2 specifically protonated PhTET2 sample demonstrated that, even in such a large system, it is still possible to detect more than 95% of the expected correlations between methyl resonances and isoleucine C β and C α resonances. These data demonstrate that this assignment strategy is also applicable to proteins above 100 kDa, provided that sequential backbone assignment is available.

We have developed a strategy for the incorporation of U-[$^{13}\text{C}'$, $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$, $^{13}\text{C}^1\text{H}_3$ - γ_2]Ile in perdeuterated proteins that is ideally suited for NMR spectroscopy studies of large proteins and complexes. An efficient synthetic route for the preparation of specifically-labelled metabolic precursors was presented that uses commercially-available reactants. We have described how this precursor can be fully incorporated into Ile residues *in vivo* without detectable isotope scrambling. The labelling pattern of the precursor has been designed to ensure that Ile residues have a “linear” [^{13}C]-labelled spin system which allows γ_2 -methyl resonances to be correlated with backbone nuclei. The new precursor and protocol we report here, therefore, further expand the repertoire of methyl-labelling protocols and will allow even greater access to structural, functional and dynamic information in large protein assemblies.

We thank Drs Franzetti, Kay and Remington for providing clones of PhTET2 and MSG; Dr Marion for critical reading of the manuscript; the Partnership for Structural Biology for access to High-Field NMR and Isotopic Labelling platforms. JB and OH acknowledge funding from ANR (ANR-09-PIRIBio-445583), and CA from CONACYT.

Notes and references

- R. Sprangers, A. Velyvis and L. E. Kay, *Nat. Methods*, 2007, **4**, 697–703; A. Ruschak and L. E. Kay, *J. Biomol. NMR*, 2010, **46**, 75.
- K. H. Gardner and L. E. Kay, *J. Am. Chem. Soc.*, 1997, **119**, 7599.
- V. Tugarinov and L. E. Kay, *J. Biomol. NMR*, 2004, **28**, 165; R. Lichtenecker, M. L. Ludwiczek, W. Schmid and R. Konrat, *J. Am. Chem. Soc.*, 2004, **126**, 5348.
- I. Ayala, R. Sounier, N. Usé, P. Gans and J. Boisbouvier, *J. Biomol. NMR*, 2009, **43**, 111.
- I. Gelis, A. M. Bonvin, D. Keramisanou, M. Koukaki, G. Gouridis, S. Karamanou, A. Economou and C. G. Kalodimos, *Cell (Cambridge, Mass.)*, 2007, **131**, 756; M. Fischer, K. Kloiber, J. Häusler, K. Ledolter, R. Konrat and W. Schmid, *ChemBioChem*, 2007, **8**, 610.
- P. Gans, O. Hamelin, R. Sounier, I. Ayala, M. A. Dura, C. D. Amero, M. Noirclec-Savoie, B. Franzetti, M. J. Plevin and J. Boisbouvier, *Angew. Chem., Int. Ed.*, 2010, **49**, 1958.
- R. Sprangers and L. E. Kay, *Nature*, 2007, **445**, 618.
- R. Sounier, L. Blanchard, Z. Wu and J. Boisbouvier, *J. Am. Chem. Soc.*, 2007, **129**, 472.
- T. L. Religa, R. Sprangers and L. E. Kay, *Science*, 2010, **328**, 98; A. M. Ruschak, T. L. Religa, S. Breuer, S. Witt and L. E. Kay, *Nature*, 2010, **467**, 868; L. E. Kay, *J. Magn. Reson.*, 2011, **210**, 159.
- A. M. Ruschak, A. Velyvis and L. E. Kay, *J. Biomol. NMR*, 2010, **48**, 129.
- V. Tugarinov and L. E. Kay, *J. Am. Chem. Soc.*, 2003, **125**, 13868.
- B. R. Howard, J. A. Endrizzi and S. J. Remington, *Biochemistry*, 2000, **39**, 3156.
- V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw and L. E. Kay, *J. Am. Chem. Soc.*, 2003, **125**, 10420; C. Amero, P. Schanda, M. A. Dura, I. Ayala, D. Marion, B. Franzetti, B. Brutscher and J. Boisbouvier, *J. Am. Chem. Soc.*, 2009, **131**, 3448.
- M. K. Rosen, K. H. Gardner, R. C. Willis, W. E. Parris, T. Pawson and L. E. Kay, *J. Mol. Biol.*, 1996, **263**, 627; D. Sheppard, C. Guo and V. Tugarinov, *J. Biomol. NMR*, 2009, **43**, 229; R. Otten, B. Chu, K. D. Krewulak, H. J. Vogel and F. A. Mulder, *J. Am. Chem. Soc.*, 2010, **132**, 2952.
- C. Guo and V. Tugarinov, *J. Biomol. NMR*, 2010, **46**, 127.
- X. Liao and V. Tugarinov, *J. Magn. Reson.*, 2011, **209**, 101.
- V. Tugarinov, D. R. Muhandiram, A. Ayed and L. E. Kay, *J. Am. Chem. Soc.*, 2003, **124**, 10025.
- M. A. Durá, V. Receveur-Brechot, J. P. Andrieu, C. Ebel, G. Schoehn, A. Roussel and B. Franzetti, *Biochemistry*, 2005, **44**, 3477; L. Borissenko and M. Groll, *J. Mol. Biol.*, 2005, **346**, 1207.