Protein Labeling

DOI: 10.1002/anie.200905660

Stereospecific Isotopic Labeling of Methyl Groups for NMR Spectroscopic Studies of High-Molecular-Weight Proteins**

Pierre Gans, Olivier Hamelin, Remy Sounier, Isabel Ayala, M. Asunción Durá, Carlos D. Amero, Marjolaine Noirclerc-Savoye, Bruno Franzetti, Michael J. Plevin, and Jérôme Boisbouvier*

Progress in NMR spectroscopy of high-molecular-weight proteins^[1] in the last decade has been strongly linked to the development of new isotopic-labeling strategies. A combination of the selective protonation of methyl groups in fully perdeuterated proteins^[2] with methyl transverse relaxation optimized spectroscopy (methyl-TROSY^[3]) has enabled local structure and dynamics to be studied in protein assemblies of up to 1 MDa in size by solution NMR spectroscopic techniques. Labeling procedures in which specifically [¹H,¹³C]methyl-labeled biosynthetic precursors are added as the sole proton source in a perdeuterated culture medium can provide a high level of methyl protonation without detectable isotopic scrambling.^[4] Maximal deuteration is critical for the optimization of resolution and the intensity of [¹H,¹³C]methyl signals in methyl-TROSY experiments.^[2e]

In initial methyl-labeling procedures, α -keto acids were used as precursors in the production of methyl-protonated isoleucine. [4a] Valine (Val) and leucine (Leu) are two amino acids of great interest, as their prochiral methyl groups account for more than 50% of all methyl probes available in proteins. [5] Leu and Val methyl groups in perdeuterated proteins are commonly protonated by using labeled α -ketoisovalerate (or 2-oxo-3-methylbutanoic acid) in which both methyl groups are ${}^{1}H$, ${}^{13}C$ -labeled. [4b] The use of this type

[*] Dr. P. Gans, Dr. R. Sounier, [+] I. Ayala, Dr. M. A. Durá, Dr. C. D. Amero, Dr. M. Noirclerc-Savoye, Dr. B. Franzetti, Dr. M. J. Plevin, Dr. J. Rojebouwjer

Institut de Biologie Structurale Jean-Pierre Ebel, CEA/CNRS/UJF 41 rue Jules Horowitz, 38027 Grenoble Cedex (France) E-mail: jerome.boisbouvier@ibs.fr

Homepage: http://www.ibs.fr/laboratories/nuclear-magneticresonance-lab/

Dr. O. Hamelin

Laboratoire de Chimie et Biologie des Métaux, Grenoble (France) I. Ayala, Dr. M. A. Durá, Dr. C. D. Amero, Dr. B. Franzetti,

Dr. M. J. Plevin, Dr. J. Boisbouvier

Centre National de la Recherche Scientifique (CNRS) (France)

Dr. P. Gans, Dr. M. Noirclerc-Savoye

Commissariat à l'Énergie Atomique (CEA) (France)

Dr. O. Hamelin

Université Joseph Fourier, Grenoble (France)

- [*] Present address: BCMP, Harvard Medical School, Boston (USA)
- [**] We thank Dr. Brutscher, Dr. Marion, and Dr. Simorre for critical reading of the manuscript, J. P. Andrieu for amino acid analysis, Prof. S. J. Remington and L. E. Kay for providing the clone of MSG, and the Partnership for Structural Biology for access to integrated structural-biology platforms. This research was supported by ANR and CNRS grants.



Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200905660.

of α -ketoisovalerate precursor proved inefficient for the study of large protein assemblies, as the intense intraresidue ¹H–¹H dipolar interaction between prochiral methyl groups limited the sensitivity of NMR spectroscopic experiments. The introduction of α-ketoisovalerate protonated at a single methyl group has been shown to significantly enhance both the resolution and sensitivity of methyl-TROSY spectra of large proteins, despite the 50% reduction in NMR-visible prochiral groups. [6] However, commercially available α -ketoisovalerate protonated at a single methyl group is a racemic mixture of pro-S and pro-R compounds, and therefore two signals are still detected for each Leu and Val residue in methyl-TROSY spectra. For high-molecular-weight proteins, this labeling strategy can result in overcrowded [¹H, ¹³C]correlation spectra owing to the sheer number of NMRvisible methyl probes. Such spectral overlap can greatly complicate the measurement of site-specific structural or relaxation parameters.

Herein we introduce a synthetic route for the production of specifically methyl-labeled acetolactate (or 2-hydroxy-2-[\frac{13}{2}]methyl-3-oxo-4-[\frac{2}{13}]butanoic acid) and demonstrate that this precursor enables the stereospecific protonation of Leu and Val methyl groups in recombinant perdeuterated proteins. Our labeling scheme offers significant enhancement of the sensitivity of [\frac{13}{2}C, \frac{1}{14}]methyl-TROSY spectra, together with a simplification of the spectra, and extends the capacity for the detection of long-range, structurally meaningful distance restraints in large proteins. As demonstrated herein, this new method will prove particularly useful for the investigation of molecular interactions within large biomolecular assemblies.

Our strategy is based on the stereospecific rearrangement of methyl groups in (S)-2-acetolactate that occurs in vivo in the early steps of Leu and Val biogenesis (Figure 1). To produce a specifically methyl-labeled acetolactate precursor, we treated unlabeled ethyl 3-oxobutanoate (1) with potassium bicarbonate in the presence of (¹³C)iodomethane. The resulting alkylated product was oxidized to ethyl 2-hydroxy-2-[13C]methyl-3-oxobutanoate (2) with molecular oxygen in the presence of cobalt(II) ions as a catalyst. 2-[13C]Methyl-4-[²H₃]acetolactate (3) was then produced by simultaneous saponification and exchange of the 4-methyl hydrogen atoms for deuterium in D₂O at pH 13.^[7] The final compound was added, without further purification, to a D2O-based Escherichia coli M9 culture medium. The specifically labeled 2S stereoisomer of acetolactate was converted in vivo into αketoisovalerate (4) by ketol-acid reductoisomerase (EC 1.1.1.86) and dihydroxyacid dehydratase (EC 4.2.1.9), [8] and then incorporated into Leu and Val. The Leu/Val biosynthetic



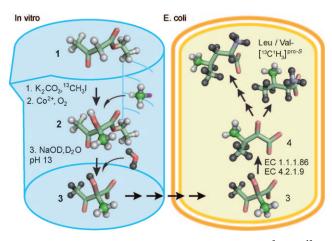


Figure 1. Schematic representation of the production of U-[²H], U-[¹²C], Leu/Val-[13C1H3]pro-S proteins. More detailed information on the synthesis is provided in Figure S1 of the Supporting Information. Compound 1: unlabeled ethyl 3-oxobutanoate; 2: ethyl 2-hydroxy-2- $[^{13}C]$ methyl-3-oxobutanoate; **3**: 2- $[^{13}C]$ methyl-4- $[^{2}H_{3}]$ acetolactate (or 2hydroxy-2-[¹³C]methyl-3-oxo-4-[²H₃]butanoic acid); **4**: (S)-3-[²H₃]methyl- $3-[^{2}H]-4-[^{13}C]-\alpha$ -ketoisovalerate (or (S)-2-oxo-3- $[^{2}H_{3}]$ methyl-3- $[^{2}H]-4$ -[13C]butanoic acid). Products 2 and 3 were produced as a racemic mixture, but only the 2S stereoisomer of the acetolactate 3 is a substrate of ketol–acid reductoisomerase (EC 1.1.1.86). 13 C, 2 H, and ¹H atoms are represented by green, black, and white balls, respectively.

pathways were saturated by adding the specifically protonated precursor at a concentration of 300 mg per liter of perdeuterated culture medium one hour prior to induction of protein overexpression.

Our new labeling strategy was applied to malate synthase G (MSG), the largest monomeric protein (82 kDa) for which a three-dimensional (3D) structure has been determined by NMR spectroscopy.^[9] Figure 2 shows two-dimensional (2D) methyl-TROSY spectra of specifically protonated MSG overexpressed in E. coli in the presence of specifically methyl-labeled α -ketoisovalerate^[6] or acetolactate 3. A direct comparison of these spectra shows that the use of the acetolactate in place of α -ketoisovalerate leads to a decrease by a factor of two in the number of correlations detected. whereas the sensitivity of the NMR spectra is increased by a factor of 1.6 ± 0.4 . This value is in agreement with the improvement in ¹³C labeling of pro-S sites (from 50 to 95%) through the use of the acetolactate (300 mg L⁻¹; see Figure S3 in the Supporting Information). The deviation from site to site reflects differences between the two labeling schemes in the local proton density for each methyl group. Only signals for Val- γ_2 and Leu- δ_2 (pro-S) methyl groups were observed when the protein was prepared by using synthetic 2-[13C]methyl-4- $[^{2}H_{3}]$ acetolactate (3). The absence of signals for Leu- δ_{1} and Val- γ_1 (pro-R) methyl groups confirms that the 13 C 1 H $_3$ methyl substituent at the 2-position of (S)-2-acetolactate is transferred stereospecifically in vivo to the pro-S position of 2,3dihydroxyisovalerate by ketol isomerase, [8] and that no methyl interconversion occurs in the following steps of the Leu/Val biosynthetic pathway. Furthermore, no signals for CH₂D and CD₂H isotopomers were detected (Figure 2B). We can therefore conclude that H/D exchange does not occur after the addition of the acetolactate to the M9/D₂O culture medium. Lastly, no scrambling to other methine, methylene, or methyl sites was detected, which indicates that the excess acetolactate in vivo does not interfere with other metabolic pathways. Thus, the use of specifically labeled acetolactate appears to be an efficient method for the complete and stereospecific methyl labeling of Leu and Val side chains in a fully perdeuterated protein background without detectable scrambling. As acetolactate is not incorporated into other metabolic pathways, it can be used conveniently in combination with other precursors, such as those proposed for the specific labeling of $\text{Ile-}\delta_1^{[4a]}$ or $\text{Ala-}\beta^{[2c]}$ methyl groups (see the Supporting Information).

An obvious first application of this new labeling scheme was the stereospecific assignment of the methyl groups of Leu and Val. Despite the existence of efficient methods to connect methyl resonances to sequentially assigned backbone nuclei,[10] the stereospecific assignment of prochiral methyl groups remains difficult. Previous approaches have involved a fractional in vivo ¹³C-labeling strategy, the use of a stereospecifically labeled library of amino acids for in vitro protein synthesis, or the measurement of small scalar couplings.^[11] Although these approaches have been shown to be useful for small and medium-sized proteins, their application to larger proteins is challenging.^[12] However, the direct stereospecific assignment of Leu and Val methyl groups is possible by visually inspecting 2D methyl-TROSY spectra of specifically protonated proteins produced with the acetolactate 3. This strategy is directly applicable to large proteins. For example, unambiguous stereospecific assignment of 98% of the Leu/ Val methyl groups of MSG (82 kDa) was possible through a simple analysis of the 2D spectrum presented in Figure 2 (see Table S1 in the Supporting Information). In this instance, sequence-specific assignments of the Leu and Val methyl groups were available.[10] When such assignments are not available, proteins can be overexpressed in the presence of uniformly ¹³C-labeled acetolactate specifically protonated at the 2-position. This labeling scheme enables pro-S methyl and backbone nuclei to be directly connected by NMR spectroscopy.^[10] Uniformly ¹³C-labeled acetolactate can be directly synthesized from commercially available uniformly 13Clabeled ethyl 3-oxobutanoate (1) by the procedure described herein.

The specific protonation of methyl groups is a powerful method for the extraction of long-range distance restraints in proteins.^[13] For large proteins, the extraction of inter-methyl NOEs is generally hampered by the low intrinsic resolution of ¹³C-edited 4D NOESY spectra. The simplification of spectra by the stereospecific labeling of prochiral methyl groups significantly reduces ambiguity in the analysis of NOE crosspeaks. NOEs between pro-S and pro-R methyl groups (Figure 3, blue cross-peaks) were no longer detected when a stereospecifically labeled sample was used. However, the replacement of $3-[^{2}H_{3}]$ methyl- $3-[^{2}H]-4-[^{13}C]-\alpha$ -ketoisovalerate^[6] (containing 50% nonstereospecific methyl protonation) with 2-[13C]methyl-4-[2H3]acetolactate increased the protonation level of pro-S methyl groups twofold, which in turn enhanced the intensity of the remaining NOE cross-peaks by a factor of four (Figure 3, black cross-peaks). This gain in sensitivity led to the detection of new, structurally meaningful

1959

Communications

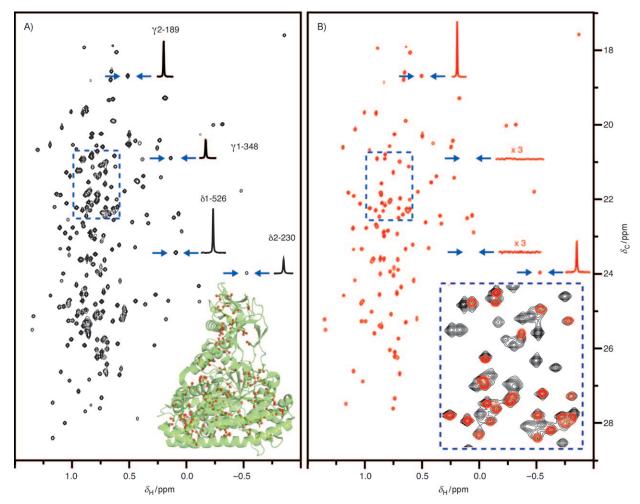


Figure 2. Comparison of methyl-TROSY spectra of MSG samples with specifically labeled Leu/Val methyl groups: $^{[10]}$ A) U- $^{[2}$ H], U- $^{[12}$ C], Leu/Val- $^{[12}$ C²H₃, 13 C¹H₃] MSG with nonstereospecific methyl labeling prepared by using 3- $^{[2]}$ H₃|methyl-3- $^{[2]}$ H₁-4- $^{[13]}$ C]-α-ketoisovalerate; $^{[6]}$ B) U- $^{[2]}$ H], U- $^{[12]}$ C], Leu/Val- $^{[13]}$ C¹H₃|p^{ro-S} MSG with stereospecific labeling prepared by using 2- $^{[13]}$ C]methyl-4- $^{[2]}$ H₃|acetolactate. Examples of 1D traces at the positions indicated by blue arrows are inserted in the figures. Correlations of Leu-δ₁ and Val-γ₁ (pro-R methyl groups) are not detected in spectrum B, whereas signals for Leu-δ₂ and Val-γ₂ (pro-S methyl groups) are enhanced by a factor of 1.6 ± 0.4. The inset in (A) shows the location of Leu and Val methyl groups (red balls) on the 3D structure of MSG. An overlay of a region of the spectra (indicated by dashed boxes) is presented in the bottom right corner of (B).

long-range NOE cross-peaks between more-remote pro-S methyl groups (Figure 3B, red cross-peaks), thereby compensating for the loss of NOEs involving pro-R methyl groups. A comparison of distance restraints extracted from 4D NOESY spectra revealed that the stereospecific labeling of Leu/Val methyl groups increased the distance threshold for which an NOE can be detected by approximately 20% (Figure 3C). The specific labeling of prochiral methyl groups is not only an attractive way to simplify the time-consuming process of NOE analysis, but it enables a significant extension of the range of structurally meaningful ¹H-¹H distances that can be measured in large proteins.

NMR spectroscopy is an established and powerful method to map biomolecular interactions with affinities ranging from the nanomolar range to the millimolar range. Nevertheless, poor resolution in NMR spectra remains a major limiting factor to the application of this method to large molecular assemblies. High-quality methyl-TROSY spectra could be obtained for a stereospecifically labeled sample of the

dodecameric aminopeptidase TET2 (468 kDa; Figure 4).^[14] Furthermore, clear changes in the spectrum were observed following the addition of amastatin, a substrate analogue of TET2 (Figure 4C). These data indicate that the labeling scheme proposed herein provides an attractive basis for the investigation of protein/ligand interaction surfaces in large biomolecular particles. Conversely, the preparation of TET2 by a nonstereospecific labeling scheme^[6] resulted in spectra with substantial cross-peak overlap, which would have greatly hampered the observation of ligand-induced chemical-shift changes (see Figure S7 in the Supporting Information).

In summary, by combining in vitro synthesis with the exploitation of naturally occurring metabolic pathways in *E. coli*, we have developed an efficient strategy for the production of perdeuterated proteins with the pro-S-specific protonation of Leu and Val methyl groups. A simple procedure based on the use of readily available reactants has been proposed for the synthesis of specifically ²H, ¹³C-labeled acetolactate, which was fully incorporated by *E. coli*

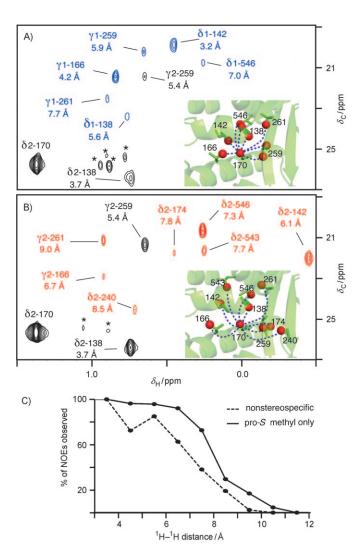


Figure 3. Observation of long-range NOEs between prochiral methyl groups in large proteins. Two-dimensional planes of 4D HMQC-NOESY-HMQC spectra of A) U- $[^{12}C]$, Leu/Val- $[^{12}C^{2}H_{3}, ^{13}C^{1}H_{3}]$ and B) U-[²H], U-[¹²C], Leu/Val-[¹³C¹H₃]^{pro-S} MSG samples. Planes were extracted at the F_1 and F_2 frequencies of the δ_2 methyl group of L170. Observable NOEs involving the δ_2 methyl group of L170 are labeled. NOE cross-peaks between pro-S and pro-R groups are displayed in blue. NOEs between pro-S methyl groups observed in both spectra are represented in black. New long-range NOEs are displayed in red. All NOEs observed in the 2D plane are shown on the 3D structure of MSG (bottom right in (A) and (B)). The shoulders of diagonal peaks located in adjacent planes are indicated by asterisks. C) Percentage of methyl-methyl NOEs observed relative to those expected, plotted as a function of the inter-methyl ¹H-¹H distance. The number of "expected NOEs" corresponds to all pairs of prochiral methyl groups separated by the distance increment with a tolerance of 0.5 Å (i.e. (6.5 ± 0.5) Å). Inter-methyl distances were derived from the X-ray crystal structure of MSG (PDB code 1d8c). Only distances between pro-S groups were considered for the stereospecifically labeled MSG sample. More detailed information on the methyl-methyl NOEs observed is provided in Figure S6 in the Supporting Information.

into the Leu and Val side chains of overexpressed proteins without scrambling to other amino acids. The stereospecific labeling of Leu/Val methyl groups decreased the number of detected correlations twofold and enhanced the intensity of

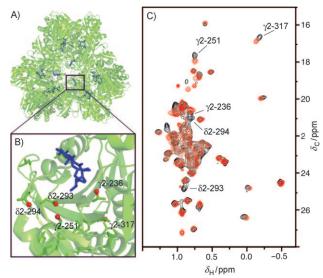


Figure 4. Interaction between TET2 (468 kDa), a peptidase involved in proteolysis in *Archaea*, and an inhibitor, amastatin (blue stick model). A) Quaternary structure of the TET2 dodecamer in a complex with amastatin (PDB code: 1Y0Y). B) Detail of the 3D structure showing Leu/Val pro-S methyl groups (red balls) affected by the binding of amastatin. C) Superposition of 2D 800 MHz methyl-TROSY spectra of U-[²H], U-[¹²C], U-[¹³C¹H₃]^{pro-S}-Leu/Val TET2, free (black) and in a complex with amastatin (red). The five labeled correlations affected by the binding of amastatin were assigned through mutagenesis.^[15]

structurally meaningful NMR signals by up to a factor of four. Furthermore, we have demonstrated that this new labeling strategy greatly simplifies the NMR spectra of protein assemblies of several hundred kilodaltons in size, and thus helps to overcome one of the current limitations preventing the study of such systems by NMR spectroscopy. We believe that the isotopic labeling scheme presented herein provides a simple and efficient way to enable solution NMR spectroscopy to be used to address to the structure and dynamics of complex biomolecular machineries and their roles in cellular processes.

Received: October 9, 2009 Revised: January 12, 2010 Published online: February 15, 2010

Keywords: isotopic labeling · methyl-group labeling · NMR spectroscopy · protein assemblies · stereospecific protonation

a) J. Fiaux, E. B. Bertelsen, A. L. Horwich, K. Wüthrich, *Nature* 2002, 418, 207–211; b) C. D. Mackereth, B. Simon, M. Sattler, *ChemBioChem* 2005, 6, 1578–1584; c) R. Sprangers, L. E. Kay, *Nature* 2007, 445, 618–622; d) R. Sprangers, A. Velyvis, L. E. Kay, *Nat. Methods* 2007, 4, 697–703.

^[2] a) V. Tugarinov, V. Kanelis, L. E. Kay, Nat. Protoc. 2006, 1, 749 – 754; b) R. L. Isaacson, P. J. Simpson, M. Liu, E. Cota, X. Zhang, P. Freemont, S. Matthews, J. Am. Chem. Soc. 2007, 129, 15428 – 15429; c) I. Ayala, R. Sounier, N. Usé, P. Gans, J. Boisbouvier, J. Biomol. NMR 2009, 43, 111 – 119; d) M. Fischer, K. Kloiber, J. Hausler, K. Ledolter, R. Konrat, W. Schmid, ChemBioChem

Communications

- **2007**, 8, 610 612; e) A. M. Ruschak, L. E. Kay, *J. Biomol. NMR* **2010**, 46, 75 87.
- [3] a) V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, L. E. Kay, J. Am. Chem. Soc. 2003, 125, 10420–10428; b) C. Amero, P. Schanda, M. A. Durá, I. Ayala, D. Marion, B. Franzetti, B. Brutscher, J. Boisbouvier, J. Am. Chem. Soc. 2009, 131, 3448–3449.
- [4] a) K. H. Gardner, L. E. Kay, J. Am. Chem. Soc. 1997, 119, 7599 7600; b) N. K. Goto, K. H. Gardner, G. A. Mueller, R. C. Willis, L. E. Kay, J. Biomol. NMR 1999, 13, 369 374.
- [5] P. McCaldon, P. Argos, Proteins Struct. Funct. Genet. 1988, 4, 99– 122.
- [6] a) V. Tugarinov, L. E. Kay, J. Biomol. NMR 2004, 28, 165–172;
 b) R. Lichtenecker, M. L. Ludwiczek, W. Schmid, R. Konrat, J. Am. Chem. Soc. 2004, 126, 5348–5349.
- [7] D. H. G. Crout, C. J. R. Hedgecock, J. Chem. Soc. Perkin Trans. 1 1979, 1982–1989.
- [8] S. R. Sylvester, C. M. Stevens, *Biochemistry* 1979, 18, 4529–4531.

- [9] V. Tugarinov, W. Y. Choy, V. Y. Orekov, L. E. Kay, *Proc. Natl. Acad. Sci. USA* 2005, 102, 622-627.
- [10] V. Tugarinov, L. E. Kay, J. Am. Chem. Soc. 2003, 125, 13868 13878
- [11] a) D. Neri, T. Szyperski, G. Otting, H. Senn, K. Wüthrich, Biochemistry 1989, 28, 7510–7516; b) G. W. Vuister, A. C. Wang, A. Bax, J. Am. Chem. Soc. 1993, 115, 5334–5335; c) M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. M. Ono, P. Güntert, Nature 2006, 440, 52–57.
- [12] V. Tugarinov, L. E. Kay, J. Am. Chem. Soc. 2004, 126, 9827–9836.
- [13] a) R. Sounier, L. Blanchard, Z. Wu, J. Boisbouvier, J. Am. Chem. Soc. 2007, 129, 472-473; b) V. Tugarinov, L. E. Kay, I. Ibraghimov, V. Y. Orekhov, J. Am. Chem. Soc. 2005, 127, 2767-2775.
- [14] a) M. A. Durá, V. Receveur-Brechot, J. Andrieu, C. Ebel, G. Schoehn, A. Roussel, B. Franzetti, *Biochemistry* 2005, 44, 3477 3486; b) L. Borissenko, M. Groll, *J. Mol Biol.* 2005, 346, 1207 1219.
- [15] R. Sprangers, A. Gribun, P. M. Hwang, W. A. Houry, L. E. Kay, Proc. Natl. Acad. Sci. USA 2005, 102, 16678–16683.