



HAL
open science

A cost-effective protocol for the parallel production of libraries of $^{13}\text{CH}_3$ -specifically labeled mutants for NMR studies of high molecular weight proteins.

Elodie Crublet, Rime Kerfah, Guillaume Mas, Marjolaine Noirclerc-Savoie, Violaine Lantez, Thierry Vernet, Jérôme Boisbouvier

► **To cite this version:**

Elodie Crublet, Rime Kerfah, Guillaume Mas, Marjolaine Noirclerc-Savoie, Violaine Lantez, et al.. A cost-effective protocol for the parallel production of libraries of $^{13}\text{CH}_3$ -specifically labeled mutants for NMR studies of high molecular weight proteins.. *Methods in Molecular Biology*, 2014, pp.229-44. hal-01093893

HAL Id: hal-01093893

<https://hal.science/hal-01093893v1>

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.

A Cost-Effective Protocol for the Parallel Production of Libraries of $^{13}\text{CH}_3$ -Specifically Labeled Mutants for NMR Studies of High Molecular Weight Proteins

Elodie Crublet, Rime Kerfah, Guillaume Mas, Marjolaine Noirclerc-Savoie, Violaine Lantez, Thierry Vernet, and Jerome Boisbouvier

Abstract

There is increasing interest in applying NMR spectroscopy to the study of large protein assemblies. Development of methyl-specific labeling protocols combined with improved NMR spectroscopy enable nowadays studies of proteins complexes up to 1 MDa. For such large complexes, the major interest lies in obtaining structural, dynamic and interaction information in solution, which requires sequence-specific resonance assignment of NMR signals. While such analysis is quite standard for small proteins, it remains one of the major bottlenecks when the size of the protein increases.

Here, we describe implementation and latest improvements of SeSAM, a fast and user-friendly approach for assignment of methyl resonances in large proteins using mutagenesis. We have improved culture medium to boost the production of methyl-specifically labeled proteins, allowing us to perform small-scale parallel production and purification of a library of $^{13}\text{CH}_3$ -specifically labeled mutants. This optimized protocol is illustrated by assignment of Alanine, Isoleucine, and Valine methyl groups of the homododecameric aminopeptidase PhTET2. We estimated that this improved method allows assignment of ca. 100 methyl cross-peaks in 2 weeks, including 4 days of NMR time and less than 2 k€ of isotopic materials.

Key words Methyl group, Isotopic labeling, High molecular weight proteins, NMR spectroscopy, SeSAM, Assignment, Site-directed mutagenesis

1 Introduction

Supramolecular systems are involved in many of the key processes that occur in cells. Therefore, understanding their local structure and dynamics is critical. For such investigations, NMR spectroscopy is a technique of choice and now allows studies of assemblies up to 1 MDa [1]. This was made possible by the development of protocols for the selective protonation of methyl groups in perdeuterated proteins [2–6]. This strategy is based on some very favorable relaxation properties of methyl groups in proteins that show

increased sensitivity compared to backbone amide proton. Moreover, methyl group containing residues are usually common and well dispersed within the polypeptide sequences, covering homogeneously the protein space. Thus, methyl groups are excellent probes of protein structure, dynamics, and interactions, particularly for very large proteins. In addition to studying naturally occurring methyl groups, methyl-containing amino acids can also be used to replace solvent-exposed residues as NMR reporters of protein interaction in order to, for instance, validate an expected binding site [7, 8].

In all cases, analysis of structural and dynamic information yielded by methyl groups requires sequence-specific assignment of methyl resonances. Conventional through bonds assignment approaches [9] to assign backbone and side chain methyl groups resonances work efficiently for small proteins but cease to be applicable to proteins over 100 kDa. In this case, alternative approaches are required. Several examples of different methyl group assignment procedures that have been successfully applied to large proteins are provided below. To date, many of the supramolecular systems studied by NMR spectroscopy are multimeric; an option to assign such assemblies is thus to try to split the quaternary complex into smaller fragments. This “divide-and-conquer” technique [1, 10] relies on disassembling the oligomeric system and transferring the resonances assignment to the full-size complex. This method however requires considerable optimization to find conditions that destabilize oligomeric interfaces without significantly disrupting the structure of the monomer or domain. Another alternative to overcome size limitation is solid-state NMR spectroscopy [11], in which the linewidth is independent of the molecular weight [12]. Yet, this approach requires crystal preparation giving high-quality spectra similar to solution state NMR, and time-consuming analysis of complex ^{13}C - ^{13}C correlation spectra acquired using solid-state NMR. Methyl-methyl Nuclear Overhauser Enhancement (NOE) experiments can also be used in combination with chemical shift prediction programs to assign methyl groups in proteins [13]. This method is so far limited to small systems due to the complexity of detecting long range NOE at more than 7 Å in very large protein assemblies. Another approach to assign large proteins is to analyze paramagnetic relaxation enhancements (PRE) induced by nitroxide spin-labels in combination with an available 3D structure [14].

Alternatively, several groups have reported a mutagenesis-based approach for assignment of some methyl resonances in large proteins. Using Leu or Val mutations along with stereospecific $^{13}\text{CH}_3$ labeling of Leu/Val residues, A. Seven and J. Rizo [15] were able to assign methyl resonances to a 73 kDa protein domain. Similarly, 15 of 17 methyl resonance frequencies of methionine methyl groups of an RNA polymerase were assigned by site-by-site

mutation strategy [16]. A similar approach was used to obtain some Ile, Leu, or Val methyl group assignments in the protease ClpB [17]. In parallel, the assignment-by-mutagenesis strategy has also been applied to a 468 kDa supramolecular protein oligomer, for the first time in a highly systematic way [18]. The method called SeSAM (Sequence-Specific Assignment of methyl groups by Mutagenesis) is based on automated molecular biology techniques, small-scale parallel preparation of residue-specific isotope-labeled samples, and sensitivity-optimized NMR experiments. Each mutant construct is expressed on a small-scale using fully perdeuterated expression media supplemented with isotope-labeled metabolic precursors designed for the specific protonation of a single class of methyl group [2–5, 10, 19–21]. A conservative mutation of one methyl-containing residue to another nonlabeled one causes the disappearance of its NMR correlation from NMR spectrum of a specifically methyl-labeled sample. This systematic strategy led to complete resonance assignment of the 34 isoleucine- δ 1 and 30 alanine- β methyl groups in less than 2 months.

Although very effective, this method remains difficult to implement. The production step is achieved from 50 mL culture medium per mutant, i.e., one culture flask for each mutant. It therefore requires a lot of manipulation from the user and may be the source of handling mistakes. Moreover, the purification step is time consuming because each mutant is purified sequentially. Therefore, we attempted to optimize and simplify this strategy (Fig. 1). First we improved the culture medium to enhance cell density, allowing us to decrease the culture volume and perform all the cultures in parallel, in 24 deep-well plates. Then, all the mutants were purified in parallel, on a 96-well plate format, therefore enabling purification of the samples in a few hours instead of weeks. Using this improved approach, we were able to reduce experiment time by a factor of 4 and isotope cost by a factor of 2 compared to previously published implementation [18].

2 Materials

2.1 Expression of Methyl-Specifically Labeled Proteins

1. Freshly transformed *E. coli* cells (BL21(DE3), BL21(DE3) RIL... etc) to overexpress protein of interest.
2. LB broth.
3. 2 \times M9 medium prepared in H₂O (for 1 L: 20 g of Na₂HPO₄, 7H₂O; 6 g of KH₂PO₄; 1 g of NaCl; 2 g of NH₄Cl). Autoclave to sterilize.
4. M9 prepared in D₂O (for 1 L: 5.3 g of anhydrous Na₂HPO₄; 3 g of anhydrous KH₂PO₄; 0.5 g of NaCl; 1 g of NH₄Cl). Use sterile D₂O.
5. Oligo-elements (for 1 L of M9 medium: 1 mL of 1 M MgSO₄, 1 mL of 0.1 M CaCl₂, 1 mL of 0.1 M MnCl₂, 1 mL of 50 mM

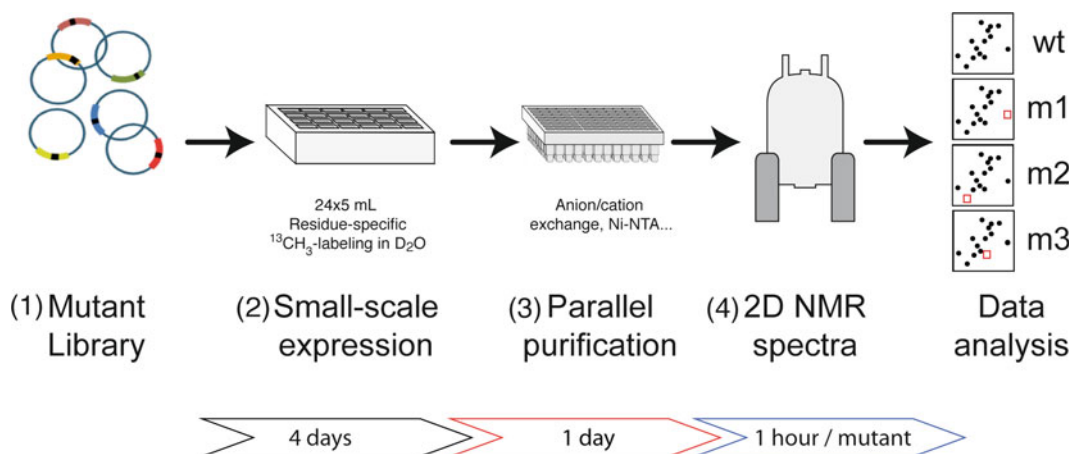


Fig. 1 The Principle of improved SeSAM. Schematic illustration of the parallel mutation-based NMR assignment strategy. (1) Each methyl-containing residue in the target sequence is mutated, on a site-by-site basis, to another similar methyl containing amino acid (e.g., Val-to-Ala) (*see* Table 1). (2) Mutant constructs are expressed on a small-scale in 24 deep-well plates, using M9 medium prepared in D₂O and supplemented with 2 g/L ²H-cell extract. One hour before induction, fully perdeuterated expression medium is complemented with isotope-labeled metabolic precursors designed for the specific incorporation of ¹³CH₃ isotopes of a single class of methyl groups. The volume of culture is adjusted to ensure a minimal yield of 0.3–0.5 mg of purified protein. (3) Cell pellets are then lysed in parallel using chemical lysis buffer and proteins are purified in 96-well plates filled with anion exchange (or any other suitable) resin. Each mutant sample is dialyzed against H₂O, lyophilized, and dissolved in NMR suitable buffer. (4) NMR spectra can be acquired using the SOFAST-methyl TROSY pulse sequence and an NMR spectrometer operating at high magnetic field. Sequence-specific assignment of each NMR signal is inferred by comparing the 2D ¹H-¹³C correlation spectrum of each member of the mutant library with a spectrum of the native protein. A conservative mutation of one methyl-containing residue to another nonlabeled one causes the disappearance of the methyl group signal from NMR spectra recorded for a specifically methyl-labeled sample

ZnSO₄, 0.5 mL of 100 mM FeCl₃). Sterilize on 0.22 μm filter. Stocks solutions should be prepared in H₂O when used in M9 100 % H₂O or 50 % H₂O/D₂O (for the precultures) and in D₂O when used in M9 100 % D₂O. In this case, all powders should be dissolved and lyophilized twice in D₂O to remove residual water before preparing stock solutions.

6. Vitamin cocktail (for 50 mL: 25 mg of pyridoxine, 25 mg of biotin, 25 mg of panthothenate hemi-calcium, 25 mg of folic acid, 25 mg of choline chloride, 25 mg of niacineamide, 2.5 mg of riboflavin, 125 mg of thiamine). Solubilize by adjusting the pH around 7, sterilize on 0.22 μm filter and decrease the pH around 5 for long-term storage. Use 2 mL for 1 L of M9 medium. Vitamins should be prepared in H₂O for precultures and in D₂O when used in M9 100 % D₂O (*see* above).
7. Isotopes: D₂O (²H ≥ 99.8 %), D-(²H, ¹²C)-glucose (²H ≥ 98 %), deuterated rich cell extract. Several sources of cell extract are commercially available (Spectra 9 (CIL), Celtone[®] Complete Medium (CIL), BioExpress[®] 1000 (CIL), Silantes[®]

E.Coli-OD2 (Silantes), Isogro® (Isotec), etc.) In this study, we chose Isogro® (noted as ^2H -cell extract in the following text), but other cell extracts are likely to give the same results.

8. IPTG (1 M in D_2O).
9. ^{13}C -methyl-specifically labeled precursors were purchased on a deuterated form ready for direct introduction into the culture medium (NMR-Bio): ^2H - ^{13}C -Alanine ($^{13}\text{C} \geq 99\%$; $^2\text{H} \geq 98\%$), ^2H - ^{13}C -2-ketobutyric acid ($^{13}\text{C} \geq 99\%$; $^2\text{H} \geq 98\%$), ^2H - ^{13}C -2-hydroxy-2-methyl-3-oxo-4-butanoic acid ($^{13}\text{C} \geq 99\%$; $^2\text{H} \geq 95\%$), ^2H -L-Isoleucine ($^2\text{H} \geq 98\%$), ^2H -L-Leucine ($^2\text{H} \geq 98\%$), ^2H - α -ketoisovalerate ($^2\text{H} \geq 98\%$).
10. 10 mL- 24 deep-well plates.
11. Gas permeable adhesive seals.

2.2 Purification

1. 10 \times BugBuster® buffer (Merck-Millipore).
2. DNase, RNase, lysozyme (Euromedex).
3. 96-well filter plates (Macherey-Nagel).
4. 2,2 mL-96 deep-well plates.
5. Aluminum seals.
6. QIAvac 96 Vacuum manifold (Qiagen) or any vacuum manifold for processing 96-well plates.
7. Appropriate resin (for nontagged proteins: any ion exchange resin (Q sepharose, SP sepharose, etc.) or affinity resin (Protein A, Protein G, Heparin sepharose, etc.); for tagged proteins: Ni-NTA, Talon resin, Strep-tactin sepharose. etc.). In this study, we used Q sepharose resin (GE Healthcare Life Sciences).
8. Anion exchange chromatography equilibration/washing buffer: 20 mM Tris-HCl pH 7.5, 160 mM NaCl.
9. Anion exchange chromatography elution buffer: 20 mM Tris-HCl pH 7.5, 350 mM NaCl.
10. Dialysis membrane (Gebaflex, Dialysis system for small-volume samples).

2.3 NMR Spectroscopy

NMR spectra were recorded on a Agilent Direct Drive spectrometer operating at a proton frequency of 800 MHz equipped with a 5 mm cryogenically cooled triple resonance pulsed field gradient probe head. Samples were loaded in a 2.5 mm shigemii tube inserted coaxially into a 5 mm tube.

3 Methods

The goal of the approach is both to improve yields and simplify the procedure previously published [18]. For that purpose, all the small-scale parallel production, lysis, and purification steps should first be set up on uniformly deuterated native protein before applying the

Table 1
Suggested substitutions for methyl containing residues

Mutated CH ₃ -containing residue	Suggested substitution	Acceptable substitution
Ala	Ser	Val [18], Gly, Cys, Thr
Ile	Val [30, 32]	Leu [17, 18, 32], Met
Leu	Ile [17], Met	Val, Phe
Val	Ile [17]	Met, Leu, Ala [25], Thr
Thr	Ser	Ala, Asn, Val
Met	Leu	Ile, Val

Adapted from BLOSUM 62 substitution matrix [29]

The BLOSUM (BLOcks SUBstitution Matrix) matrix is a substitution matrix used for sequence alignment of proteins [29]. BLOSUM matrices are used to score alignments between evolutionarily divergent protein sequences. They are based on local alignments. In this table, suggested/acceptable substitutions for each of the 6 methyl-containing amino acids, for the purpose of assignment using SeSAM strategy [18], are listed along with an application reference when it exists

protocol to the whole isotopically labeled library of mutants. The hypothesis is made that all mutants behave identically to the native protein and that methyl labeling does not change the purification profile (same expression level, same purification conditions, etc.).

3.1 Generation of Mutant Plasmids Libraries

Constructs carrying single point mutations can be purchased commercially or prepared using an automated molecular biology platform. Here, they were generated by in-house automated molecular biology Platform (RoBioMol—Institut de Biologie Structurale, Jean-Pierre Ebel) using an automated PCR-based protocol adapted from the QuikChange site-directed mutagenesis method. PCR amplification was performed with Phusion Hot Start enzyme (Finnzymes) using the expression plasmid pET4Ic-PhTET2 as template and the specific mutagenic primers. Products were purified and digested by DpnI. Final mutations were selected by transformation and verified by sequencing. Valine residues were mutated into Alanine residues, Alanine into Valine, Isoleucine into Leucine. Amino acids must be substituted by an isosteric one to avoid considerable changes in the structure and minimize secondary shift effects (*see* Note 1 and Table 1).

3.2 Protein Labeling for Methyl Detection

Selective methyl protonation provides excellent probes for monitoring interactions and dynamics, and high-quality spectra can be recorded in very large systems. For assignment, a prerequisite is to restrict labeling to only one type of methyl-containing residue at once, in order to minimize secondary chemical shift perturbations

due to the mutation. Currently, methods are available to label Isoleucine ($\delta 1$ [2] or $\gamma 2$ [22]), Alanine [3, 23], Threonine [24], or Methionine [10, 20] methyl groups efficiently. For Leucine and Valine, however, the labeling scheme is more difficult because these amino acids share the same biogenesis pathway and each residue has two prochiral methyl groups. Consequently, compared to other residues, for a similar spectral window, the number of methyl group resonances is four times higher, potentially resulting in peak overlap and loss of resolution. Thus, until recently, the assignment-by-mutagenesis strategy was not attractive for Valine residues. Since then, a stereospecific isotopic labeling method was developed for Leucine/Valine labeling [21]. More recently, our group has also developed the stereoselective labeling of Valine without Leucine using coinorporation of pro-*S* Acetolactate and ^2H -Leucine [25]. This new isotopic labeling strategy increased the resolution and reduced overlaps and potential secondary chemical shifts by a factor of 4 on average. Thus, it enabled easy assignment-by-mutagenesis of Valine residues in large protein assemblies, compared to standard labeling using 2 keto-acids [26].

In this section we describe the protocol to specifically label large proteins for methyl detection and produced them on a small-scale using enriched cell extract.

1. Bacteria must first be trained to adapt progressively in 100 % ^2H medium. All these adaptation steps will be performed in 24-well plates covered with gas permeable adhesive seal. Pick a fresh colony of transformed cells from each mutant of an LB plate (*see Note 2*) and start a 2-mL bacterial culture of LB medium prepared in H_2O . Once $\text{OD}_{600\text{nm}}$ reaches 3–4 (6–8 h) at 37 °C, transfer 50 μL of the cell culture into 2 mL of M9 medium prepared in H_2O . The starting $\text{OD}_{600\text{nm}}$ of this new culture should be about 0.1. Incubate the culture in a shaking incubator (220–250 rpm) at 37 °C overnight (the $\text{OD}_{600\text{nm}}$ should be around 2–2.5) and transfer 100–200 μL of the cell culture into 2 mL of M9 medium prepared in 50 % D_2O (*see Note 3*) (starting $\text{OD}_{600\text{nm}} = 0.2$). Let the culture grow at 37 °C until the $\text{OD}_{600\text{nm}}$ reaches 2–2.5 and transfer 300 μL of the cell culture into 3 mL of M9 medium in 100 % D_2O (starting $\text{OD}_{600\text{nm}} = 0.25$). Incubate the culture at 37 °C overnight until $\text{OD}_{600\text{nm}}$ is 1.5–2 and use this cell culture as your starting culture.
2. Prepare the volume of M9 needed for the whole culture (2×5 mL for each mutant) in a sterile flask (M9 in D_2O + oligo-elements in D_2O + vitamins in D_2O + 2 g/L of ^{12}C , ^2H glucose + antibiotics). Add 2 g/L of ^2H -cell extract (*see Notes 4–6*).
3. Fill a 24 deep-well block (*see Note 7*) with 3.5 mL of this medium. Inoculate each well with the overnight culture at a

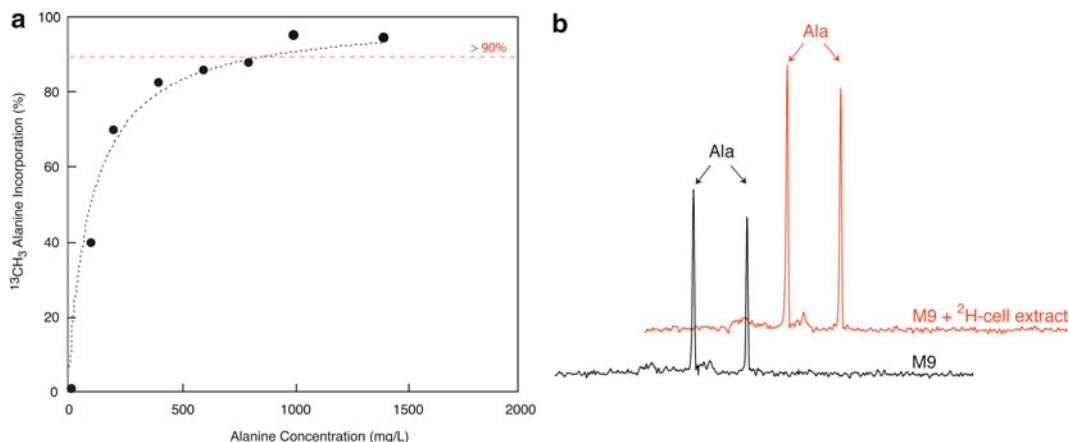


Fig. 2 Level of incorporation of $^{13}\text{CH}_3$ -alanine in proteins expressed in M9 medium supplemented with ^2H -cell extract. Level of incorporation of ^{13}C at the $\text{C}\beta$ -alanine position in overexpressed ubiquitin as a function of the amount of the exogenous alanine added in culture medium 1 h prior to induction. Ubiquitin was expressed in *E. coli* in M9/ D_2O culture medium supplemented with 2 g/L ^2H , ^{12}C glucose, and 2 g/L ^2H , ^{12}C Isogro $^{\text{®}}$. Different amounts of $^{13}\text{CH}_3$ L-alanine (0–1.4 g/L) were added 1 h before induction. ^2H -Isovalerate (400 mg/L) and ^2H -Isoleucine (120 mg/L) were added to prevent scrambling from alanine. Fixed amount of $^{13}\text{CH}_3$ L-methionine (0.5 g/L) was used as an internal reference. The level of incorporation was determined by analyzing the intensities of one alanine methyl resonance with respect to signal of methyl group of one methionine residue. A level of incorporation in Alanine side chains $\geq 90\%$ is obtained by adding 1 g of alanine per liter of culture medium. **(b)** Comparison of 1D ^{13}C -filtered NMR spectra of U- $[\text{C}^{12}, \text{H}^2]$, Ala- $[\text{C}^{13}\text{H}_3]\beta$ ubiquitin expressed in equal volume of M9 medium (M9) or M9 medium supplemented with ^2H -cell extract (M9+ ^2H -cell extract). Data were recorded at 37 °C using a 2.5 mm Shigemi tube, on an 800 MHz NMR spectrometer equipped with a cryogenic probe head. Comparison of 1D spectra shows an increase of the signal-to-noise ratio by a factor of 1.65. These results are in agreement with an increase of both the cell density ($\text{OD}_{600\text{nm}} \times 1.7$ with ^2H -cell extract) and the culture yields (yields $\times 1.75$ with ^2H -cell extract), indicating that ^{13}C -Alanine is fully incorporated in these conditions

starting $\text{OD}_{600\text{nm}}$ of 0.3. Cover with gas permeable adhesive seal and grow at 37 °C until $\text{OD}_{600\text{nm}}$ reaches 1.5.

4. Add the precursors (*see Note 8*) diluted in 1.5 mL of M9/ D_2O , 1 h prior to IPTG induction. The amount of $^{13}\text{CH}_3$ -precursors to add was optimized to ensure complete labeling of proteins overexpressed in rich medium (Fig. 2a). The yield of the protein was improved by a factor of 1.7 when ^2H -cell extract was added to the culture medium, while the frequency of labeling was still almost 100 %, proving there is no isotopic dilution by ^2H -cell extract (Fig. 2b).
5. Let the culture grow for 1 h. The $\text{OD}_{600\text{nm}}$ should reach a value of 1.5.
6. Add IPTG (in D_2O) to 0.5 mM to induce protein expression. Continue incubation at 37 °C for 4 h (depending on your protein).

7. Centrifuge the entire plate in a swing-out rotor for 20 min at $3,250\times g$. Discard the supernatant and store the pellet at $-80\text{ }^{\circ}\text{C}$ or process directly.

3.3 Lysis

Protein purification and lysis strategy may vary according to the protein of interest (presence and type of fusion tag). Here we described the lysis strategy optimized for PhTET2 expressed in BL21(DE3)RIL.

1. Lyse the cells (*see Note 9*). Add 250 μL of lysis buffer into each well in the 24-well culture plate ($1\times$ BugBuster[®], 20 mM Tris-HCl pH 7.5, 100 mM NaCl (*see Note 10*), 5 mM MgCl_2 , 2 $\mu\text{g}/\text{mL}$ DNase, 10 $\mu\text{g}/\text{mL}$ RNase, 0.3 mg/mL Lysozyme) and resuspend the cells by pipetting up and down. Pool the 2 pellets of the same mutant (*see Note 7*) and incubate 40 min at room temperature with (occasional) shaking.
2. Heat the crude extract at $85\text{ }^{\circ}\text{C}$ (*see Note 11*) for 15 min by making the plate directly float in a water bath.
3. Centrifuge the plate at $3,250\times g$ for 20 min. Both soluble- and whole-cell pellets should be analyzed (*see Subheading 3.4 step 6*).

3.4 Small-Scale Parallel Purification

In this section, we describe a typical protein purification procedure using an anion exchange resin, but alternative resins can be used (*see Note 12*).

1. Prepare the purification 96-well filter plate. Resuspend the Q sepharose resin thoroughly. Pipet 800 μL of resin suspension (bed volume of 400 μL) into each well of the plate. Wash each well twice with 1 mL of water and 3 times with 1 mL of equilibration buffer using a vacuum manifold (or alternatively a centrifuge with a swing-out rotor).
2. Transfer the clear supernatants (Subheading 3.3 step 3) into the 96-well filter plate containing the resin. Seal the block with aluminum seal and caps to avoid leaks and place at room temperature for 1 h with gentle shaking.
3. Place the 96-well filter plate over a 96 deep-well plate in the vacuum manifold. Remove the aluminum seal and caps (only over the used wells) and let the samples flow through the resin first by gravity, then by applying vacuum until the samples have been completely drawn through the plate.
4. Place a drain deep-well plate and wash the resin by adding 400 μL of washing buffer to each well and then apply vacuum as above. Repeat the wash 4 times.
5. Place the filter plate on top of a new deep-well plate and add 400 μL of elution buffer in each well. Incubate 2 min and proceed as above. Repeat the elution 4 times and store the eluate.

6. Analyze by SDS-PAGE the total, soluble and eluted fractions for some of the mutants, according to standard procedures.

3.5 NMR Spectroscopy

To assign a large protein such as PhTET2 (468 kDa, 353 residues per subunit), the SeSAM strategy [18] is followed.

The type of probe and the tube configuration can affect the required amount of material. For small sample amounts, users should always choose the most sensitive probe available (best signal-to-noise per mg of protein). Using a standard 5-mm cryogenically cooled probe, the protein should be concentrated to optimize the amount of spins present in the most sensitive area of the active volume (i.e., near the axial symmetry axis). This can be achieved simply by placing the sample in a small diameter tube (1 to 3 mm) centered inside a standard 5-mm tube. As a result, compared to a 5-mm NMR tube, the sensitivity gain is ca. a factor of 2 per mg of protein. Alternatively, using cryoprobes optimized for small volumes (3 mm coldprobe (Agilent); 1.7 mm microcryoprobe (Bruker)) can further increase sensitivity twofold allowing a reduction of the culture volume (one 5-mL well for each mutant) or a division of the acquisition time by a factor of 4. In this study, the samples were loaded in a 2.5 mm shigemi tube placed coaxially to a regular 5-mm NMR tube as a sample holder and NMR spectra were recorded on a spectrometer equipped with a 5-mm cryoprobe.

1. Prepare methyl-labeled mutants as described above. Transfer the protein in an NMR suitable buffer. In our case, the protein was extensively dialyzed in H₂O, lyophilized and resuspended in 60 μ L (\pm 0.15 mM) of 20 mM Tris-DCl, 20 mM NaCl pH 7.4 in 100 % D₂O. Alternatively (if the protein is not stable in pure H₂O), buffer can be exchanged by a series of dilution/concentration steps in D₂O buffer using a concentration unit with an appropriate cut-off.
2. Record ¹H-¹³C SOFAST-methyl-TROSY spectra (more sensitive experiment by unit of time [27]) for each mutant sample and native protein. Here, taking advantage of the thermostability of PhTET2, NMR data are recorded at 50 °C. The angle of proton excitation pulse is set to 30° and the recycling delay is optimized to 0.4 s. The length of each NMR experiment is adjusted depending on the concentration of the sample (typically 1 h for 9 nmol of sample).
3. All data are processed and analyzed with NMRPipe [28]. Sequence-specific assignments of each NMR signal are inferred by comparing the 2D ¹H-¹³C correlation spectrum recorded for each member of the mutant library with a spectrum recorded for the wild-type protein (Fig. 3). Conceptually, assignment-by-mutagenesis is straightforward. In practice, however, the overlap of resonances and the occurrence of secondary chemical

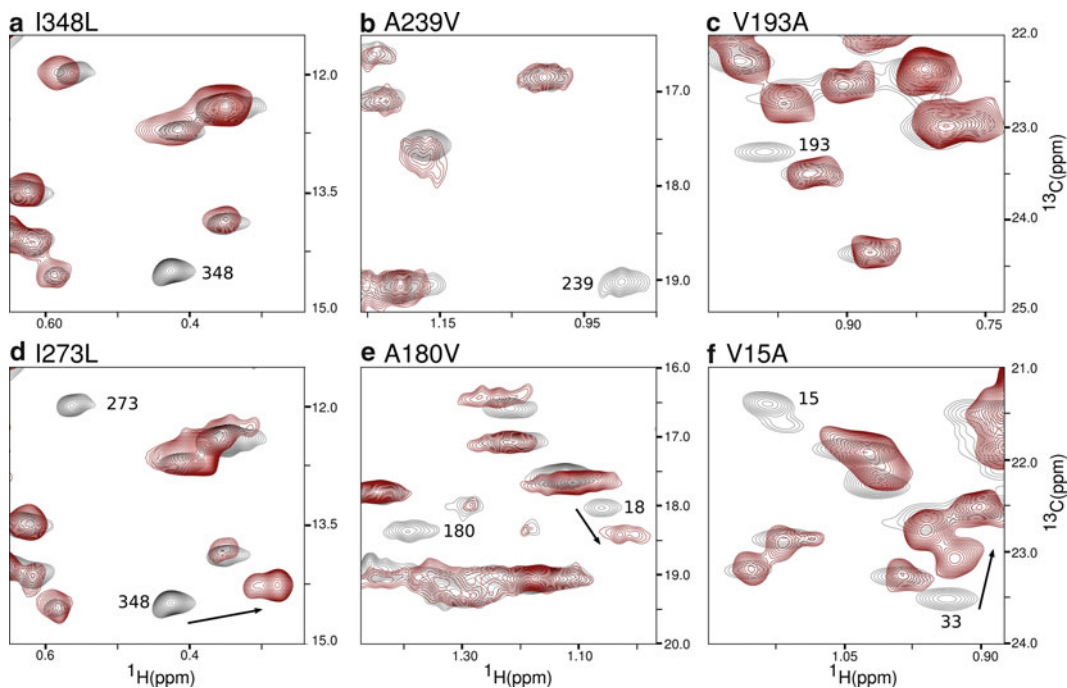


Fig. 3 Examples of spectra of mutants with specific Ile- δ 1, Ala- β , or Val pro-*S*-labeled methyl probes. Spectra of mutants displaying modest secondary chemical shift perturbations (**a–c**) were initially chosen for analysis and, when possible, sequence-specific assignment of methyl groups were made. This initial set of unambiguous assignments assisted the analysis of spectra displaying larger chemical shift perturbations (**d–f**). Spectra of this figure are extracted from the work of Amero et al. [18] and Mas et al. [25]. SOFAST Methyl-TROSY spectra were recorded at 50 °C, on an 800 MHz NMR spectrometer equipped with a cryogenic probe head, using samples of U- ^{12}C , ^2H , Ile- $^{13}\text{CH}_3$ δ^1 (**a** and **d**) [18], U- ^{12}C , ^2H , Ala- $^{13}\text{CH}_3$ β (**b** and **e**) [18], or U- ^{12}C , ^2H , Val- $^{13}\text{CH}_3$ $^{\text{pro-}S}$ (**c** and **f**) [25] labeled mutant PhTET2 protein (*red*). Each spectrum was overlaid with the reference spectrum of the native particle (*black*). Experimental acquisition times were adjusted to the sample protein concentration (from 0.26 to 0.42 mg/sample, i.e., 7–11 nmol of PhTET2 monomer) with a maximum experimental time set to 1 h. The assignment inferred for the missing resonance in the mutant spectrum is indicated and secondary chemical shift perturbations are annotated with *arrows*

shifts changes can confuse the analysis. Perturbations are likely to occur but can be minimized (using conservative mutations). In the first round, all spectra with only one missing peak are considered for a straightforward assignment of a first set of resonances. Then, more complex spectra are studied with consideration of structure, first set of unambiguous assignment and the whole set of spectra taken into account, in order to analyze secondary chemical shift perturbations. That is why considering the full library of single-site mutations greatly simplifies the process of resonance assignment by cross-validating the results several times (*see Note 13*).

4 Conclusion

Thanks to the development of protocols and molecules that allow residue-specific protonation of methyl groups in highly perdeuterated proteins, it is now feasible to apply solution NMR techniques to protein systems as large as 1 MDa. Structural and dynamic information yielded by methyl groups is most useful when a sequence-specific assignment for the probe is known. Easily obtaining these assignments remains the major hurdle in many studies of large proteins. Here we describe a fast, efficient, and user-friendly protocol for resonance assignment that has allowed us to assign up to 100 methyl cross-peaks in about 2 weeks with 4 days of NMR time and a isotopic cost of less than 2 k€. We demonstrated the feasibility of this protocol on samples labeled on Alanine, Isoleucine, or Valine. This method can also be extended to Methionine- or Threonine-labeled proteins. As for Valine labeling, the methyl-specific labeling of Leucine residues will require the development of a stereoselectively labeled amino acid (or precursor) to label Leucine residues independently from Valines.

5 Notes

1. Amino acids should typically be substituted by an isosteric amino acid to prevent significant changes in the local environment and protein packing, which could introduce significant chemical shift perturbations. We chose to exchange Valine to Alanine according to Amero et al. [18], who mutated Alanine into Valine. However, according to BLOSUM matrices for amino acid substitutions [29], the Valine to Isoleucine mutations (score 3) would have been wiser and most likely to have minimal effect on the structure and NMR spectra. Indeed, Valine differs from Isoleucine by the loss of only one CH₂ group, whereas it differs from Alanine by the loss of CH-CH₃. According to these matrices, Chan et al. [30] generated a set of mutants in which each Isoleucine was substituted by a Valine (score 3). The resulting chemical shift perturbations were smaller than those from the corresponding Ile to mutations Ala (score -1) facilitating the assignment (Table 1).
2. Antibiotics are not specified. Put appropriate antibiotics considering the plasmid and the cells used. However, ampicillin is not suitable because it can be inactivated by the β -lactamases produced by the cell, resulting in plasmid loss and drop of protein expression. This effect seems to be worse in D₂O medium because of the successive adaptation steps that promote loss of selective pressure. Here, we routinely used the pET-41 plasmid, which contains the kanamycin resistance gene.

3. M9 50 % D₂O is prepared with 2× M9 H₂O and same volume of D₂O. The oligo-elements, vitamins, antibiotics, and glucose used at this step are still prepared in water.
4. To increase cell density, we supplemented M9 medium with ²H-cell extract. We optimized the protocol with Isogro[®] but many similar rich bacterial cell growth media are commercially available (Spectra 9, Celtone[®] Complete Medium, BioExpress[®] 1000, Silantes[®] E.Coli-OD2, etc.). We did not test them but they can most likely be used to replace Isogro[®], after suitable optimization.
5. Rich medium can be the source of isotopic dilution as it contains ¹²C²H₃-labeled amino acids. To minimize this, we tested different concentrations of ²H-cell extract (2, 3, or 10 g/L). Growth is significantly increased from 0 to 2 g/L; however higher concentrations of ²H-cell extract did not show further significant improvement. Moreover, ²H-cell extract is added at the start of the culture, whereas methyl specifically labeled precursors are added 1 h before induction. This early addition reduces isotopic dilution because most of deuterated amino acids are consumed for the cell growth, while ¹³CH₃-precursors are incorporated for the production of the labeled protein. Optimization was done using Isogro[®], but similar results are expected for other types of commercial ²H-cell extracts.
6. For ubiquitin, yields with 2 g/L of ²H-cell extract were increased by a factor of 1.7 (Fig. 2b).
7. Considering both yields, parallelization of the method and simplification of the purification steps, we were able to reduce the volume of production to 10 mL for each mutant. One 24 deep-well plate is thus suitable for 12 mutants (2 wells of 5 mL for each mutant). The final concentration in a 2.5 mm Shigemi tube is around 0.15 mM (9 nmol or 0.35 mg of PhTET2 monomer) in 60 μL.
8. Incorporation level of ¹³CH₃-alanine was determined using 2 g/L of ¹²C, ²H-glucose, and 2 g/L of ²H-cell extract in culture medium. Using higher concentrations of glucose may cause isotopic dilution of the ¹³CH₃-precursor [31] and the incorporation curves must then be modified accordingly.
9. One crucial step to optimize in the protein production process is the bacterial cell lysis. Different lysis methods should be tested. However, when tens or hundreds of mutants are produced, only a few lysis methods can be reasonably used in parallel. Sonication proved problematic unless using High Throughput sonicators (which are expensive) or ultrasonic bath (but the ultrasonic energy is not always equally distributed throughout the plate and the results may not be reproducible from well to well). For these reasons, we chose to

optimize lysis conditions using chemical treatment (all the preliminary tests were performed on the native protein). This allowed us to achieve lysis directly on the culture plates, in parallel. We tested different lysis buffers and selected the BugBuster[®], as it showed the best solubility yields for our protein of interest.

10. The NaCl concentration of the lysis buffer has to be settled according to the protein of interest. Decrease it if the protein does not bind to the ion exchange resin.
11. The protein studied is an aminopeptidase from a thermophilic organism, adapted to high temperatures. The protein is thus stable at 85 °C for 15 min, whereas most of other *E. coli* proteins will precipitate. Bacterial contaminants are then removed by centrifugation.
12. Fill your filter plate with any suitable resin (anion or cation exchange resin, Ni-NTA, StrepTactin, etc.). Adapt bead volume to the capacity of the resin. Typically, we prepare wells containing 400 μL of anion exchange resin for a 10-mL cell culture. This can be scaled up or down to suit your needs based on the expected protein yield.
13. In almost 50 % of cases (using conservative mutations (*see Note 1*)), the only difference between mutant and reference spectra is a single missing cross-peak in the spectrum of the mutant (Fig. 3a–c). In such instances, the missing peak can be unambiguously assigned to the methyl group of the mutated residue. In the remaining spectra, the disappearance of the signal is accompanied by small changes in the chemical shift of a few additional correlations (Fig. 3d–f). This effect is expected and has previously been observed [18, 32]. Peak movements that do not directly concern the mutated resonance can complicate the process of obtaining a sequence-specific assignment from a single experiment, especially in an overcrowded region of the spectrum. Conservative mutations enable minimization of secondary chemical shift perturbations. In the same way, stereospecific labeling of a single methyl group (Valine pro-S) versus labeling of 4 methyl groups (Val/Leu) using α -ketoisovalerate [26] reduces peak overlapping as well as secondary chemical shift up to a factor of 4.

Nonetheless, secondary chemical shift perturbations reflect modifications in the local electronic environment and can therefore provide complementary information that can be used to confirm the proposed assignment. The key point is that the information provided by secondary chemical shift changes only becomes interpretable when data from a full library of methyl group mutants is considered [18]. Any ambiguous assignment can therefore be readily cross-validated

using structurally close, straightforward-assigned resonances. Using an incomplete library of mutants would not permit the same level of confidence in the final assignments.

Acknowledgments

We would like to thank Dr P. Macek, M. Plevin, O. Hamelin, P. Gans, I. Ayala, C. Amero, and A. Favier for stimulating discussions and assistance in sample preparation or analysis. This work used the RobioMol, High-Field NMR, Isotopic Labeling, and Seq3A platforms of the Grenoble Instruct centre (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). The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme FP7/2007-2013 Grant Agreement no. 260887.

References

1. Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature* 445:618–622
2. Gardner K, Kay LE (1997) Production and incorporation of ^{15}N , ^{13}C , ^2H (^1H - δ 1 methyl) isoleucine into proteins for multidimensional NMR studies. *J Am Chem Soc* 119:7599–7600
3. Ayala I, Sounier R, Use N et al (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. *J Biomol NMR* 43:111–119
4. Tugarinov V, Kanelis V, Kay LE (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat Protoc* 1:749–754
5. Ruschak AM, Kay LE (2010) Methyl groups as probes of supra-molecular structure, dynamics and function. *J Biomol NMR* 46:75–87
6. Plevin MJ, Boisbouvier J (2012) Isotope-labelling of methyl groups for NMR studies of large proteins. In: *Recent developments in biomolecular NMR*. Royal Society of Chemistry. doi:10.1039/9781849735391
7. Stoffregen MC, Schwer MM, Renschler FA et al (2012) Methionine scanning as an NMR tool for detecting and analyzing biomolecular interaction surfaces. *Structure* 20:573–581
8. Religa TL, Ruschak AM, Rosenzweig R et al (2011) Site-directed methyl group labeling as an NMR probe of structure and dynamics in supramolecular protein systems: applications to the proteasome and to the ClpP protease. *J Am Chem Soc* 133:9063–9068
9. Bax A (2011) Triple resonance three-dimensional protein NMR: before it became a black box. *J Magn Reson* 213:442–445
10. Gelis I, Bonvin AM, Keramisanou D et al (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 131:756–769
11. Turano P, Lalli D, Felli IC et al (2010) NMR reveals pathway for ferritin mineral precursors to the central cavity of ferritin. *Proc Natl Acad Sci USA* 107:545–550
12. Marassi FM, Ramamoorthy A, Opella SJ (1997) Complete resolution of the solid-state NMR spectrum of a uniformly ^{15}N -labeled membrane protein in phospholipid bilayers. *Proc Natl Acad Sci USA* 94:8551–8556
13. Xu Y, Liu M, Simpson PJ et al (2009) Automated assignment in selectively methyl-labelled proteins. *J Am Chem Soc* 131:9480–9481
14. Venditti V, Fawzi NL, Clore GM (2011) Automated sequence- and stereo-specific assignment of methyl-labeled proteins by para-

- magnetic relaxation and methyl-methyl nuclear overhauser enhancement spectroscopy. *J Biomol NMR* 51:319–328
15. Seven A, Rizo J (2012) Assigning the methyl resonances of the 73 kDa Munc13-1 MUN domain by mutagenesis. 25 th ICMRBS-poster n° P302 TU, Lyon
 16. Yang X, Welch JL, Arnold JJ et al (2010) Long-range interaction networks in the function and fidelity of poliovirus RNA-dependent RNA polymerase studied by nuclear magnetic resonance. *Biochemistry* 49:9361–9371
 17. Rosenzweig R, Moradi S, Zarrine-Afsar A et al (2013) Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science* 339:1080–1083
 18. Amero C, Asuncion Dura M, Noirclerc-Savoye M et al (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J Biomol NMR* 50:229–236
 19. Goto NK, Gardner KH, Mueller GA et al (1999) A robust and cost-effective method for the production of Val, Leu, Ile (δ 1) methyl-protonated ^{15}N -, ^{13}C -, ^2H -labeled proteins. *J Biomol NMR* 13:369–374
 20. Fischer M, Kloiber K, Hausler J et al (2007) Synthesis of a ^{13}C -methyl-group-labeled methionine precursor as a useful tool for simplifying protein structural analysis by NMR spectroscopy. *Chembiochem* 8:610–612
 21. Gans P, Hamelin O, Sounier R et al (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chem Int Ed Engl* 49:1958–1962
 22. Ayala I, Hamelin O, Amero C et al (2012) An optimized isotopic labelling strategy of isoleucine-gamma2 methyl groups for solution NMR studies of high molecular weight proteins. *Chem Commun (Camb)* 48:1434–1436
 23. Isaacson RL, Simpson PJ, Liu M et al (2007) A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues. *J Am Chem Soc* 129:15428–15429
 24. Sinha K, Jen-Jacobson L, Rule GS (2011) Specific labeling of threonine methyl groups for NMR studies of protein-nucleic acid complexes. *Biochemistry* 50:10189–10191
 25. Mas G, Crublet E, Hamelin O et al (2013) Specific labeling and assignment Strategies of valine methyl groups for the NMR Studies of high molecular weight proteins (submitted)
 26. Tugarinov V, Kay LE (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J Am Chem Soc* 125:13868–13878
 27. Amero C, Schanda P, Dura MA et al (2009) Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies. *J Am Chem Soc* 131:3448–3449
 28. Delaglio F, Grzesiek S, Vuister GW et al (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293
 29. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89:10915–10919
 30. Chan PH, Weissbach S, Okon M et al (2012) Nuclear magnetic resonance spectral assignments of α -1,4-galactosyltransferase LgtC from *Neisseria meningitidis*: substrate binding and multiple conformational states. *Biochemistry* 51:8278–8292
 31. Plevin MJ, Hamelin O, Boisbouvier J et al (2011) A simple biosynthetic method for stereospecific resonance assignment of prochiral methyl groups in proteins. *J Biomol NMR* 49:61–67
 32. Sprangers R, Gribun A, Hwang PM et al (2005) Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. *Proc Natl Acad Sci USA* 102:16678–16683