J Biomol NMR (2016) 65:15–27
DOI 10.1007/s10858-016-0032-2

ARTICLE

Facilitating unambiguous NMR assignments and enabling higher probe density through selective labeling of all methyl containing amino acids

Andrew Proudfoot[1] · Andreas O. Frank[1] · Fiorella Ruggiu[1] · Mulugeta Mamo[1] · Andreas Lingel[1]

Received: 30 January 2016 / Accepted: 19 April 2016 / Published online: 29 April 2016

Abstract The deuteration of proteins and selective labeling of side chain methyl groups has greatly enhanced the molecular weight range of proteins and protein complexes which can be studied using solution NMR spectroscopy. Protocols for the selective labeling of all six methyl group containing amino acids individually are available, however to date, only a maximum of five amino acids have been labeled simultaneously. Here, we describe a new methodology for the simultaneous, selective labeling of all six methyl containing amino acids using the 115 kDa homohexameric enzyme CoaD from *E. coli* as a model system. The utility of the labeling protocol is demonstrated by efficiently and unambiguously assigning all methyl groups in the enzymatic active site using a single 4D ^{13}C-resolved HMQC–NOESY–HMQC experiment, in conjunction with a crystal structure. Furthermore, the six fold labeled protein was employed to characterize the interaction between the substrate analogue (R)-pantetheine and CoaD by chemical shift perturbations, demonstrating the benefit of the increased probe density.

Keywords Selective labeling · 4D NOESY · CoaD · Methyl groups · Assignments · NMR

Electronic supplementary material The online version of this article (doi:10.1007/s10858-016-0032-2) contains supplementary material, which is available to authorized users.

✉ Andreas Lingel
andreas.lingel@novartis.com

[1] Novartis Institutes for BioMedical Research, 5300 Chiron Way, Emeryville, CA 94608, USA

Introduction

Since the mid-1990s, NMR methodologies have been developed to enable the study of larger proteins and complexes up to 1 MDa in size (Gardner and Kay 1997; Sprangers and Kay 2007). Improvements in technology and methodology have come in the forms of improved hardware (i.e. higher field NMR spectrometers along with cryogenic probes), optimized NMR experiments (i.e. TROSY) (Yamazaki et al. 1994; Pervushin et al. 1997; Tugarinov et al. 2003) as well as the development of labeling strategies to perdeuterate proteins and to re-introduce NMR-active nuclei in methyl groups at specific locations on the amino acid side chain (Goto and Kay 2000; Tugarinov et al. 2004; Religa and Kay 2010; Kerfah et al. 2015c). Gardner, Goto and Kay pioneered the selective labeling of amino acids in the 1990's when they published work detailing how to selectively label the δ_1-methyl position of isoleucine and the δ-methyl positions of leucine as well as the γ-methyl positions of valine by supplementing the growth medium with [3,3-^{2}H$_2$] ^{13}C 2-ketobutyrate and [3-^{2}H] α-ketoisovalerate, respectively (Gardner and Kay 1997; Goto et al. 1999). Since then, labeling schemes have been expanded to the extent that all six methyl containing amino acids can be individually labeled in separate samples either through the use of labeled amino acid precursors (I, L, V) (Gardner and Kay 1997; Goto et al. 1999; Hajduk et al. 2000; Gross et al. 2003), or by supplementing the growth medium directly with selectively labeled amino acids (M, A, T) (Isaacson et al. 2007; Fischer et al. 2007; Gelis et al. 2007; Ayala et al. 2009; Stoffregen et al. 2012; Velyvis et al. 2012). In addition to being able to selectively label individual residues, protocols have been developed for the simultaneous labeling of multiple amino acids including ILV, MILV,

ILVA or MILVT (Gross et al. 2003; Tugarinov and Kay 2003b; Lichtenecker et al. 2004; Gelis et al. 2007; Velyvis et al. 2009; Godoy-Ruiz et al. 2010; Saio et al. 2014). However, to the authors' knowledge, to date no protocol has been published where all six methyl containing amino acids have been labeled simultaneously.

In order to ensure that detailed structural and functional information can be obtained from NMR experiments, resonance assignment is required. The use of point mutations to assign methyl groups in selectively labeled proteins has been well described and has become common practice (Sprangers and Kay 2007; Velyvis et al. 2009; Religa et al. 2010). However, this process can be fairly labor and resource intensive and mutations frequently result in significantly reduced or limited protein expression. In cases where high protein expression is obtained, replacement of a methyl group bearing amino acid with a similar residue (Crublet et al. 2014) can lead to dramatic changes in the chemical shifts of amino acids in spatial proximity to the mutation site, making unambiguous identification of the missing methyl peak challenging (Xu and Matthews 2013a). Due to these difficulties, a number of additional methods have been developed for the assignment of methyl groups, including the implementation of new pulse programs which correlate the methyl group of certain amino acids to the backbone via the aliphatic side chain (Tugarinov and Kay 2003a, b; Kerfah et al. 2015a), or alternatively the use of high dimensional methyl–methyl NOESY spectra to utilize methyl–methyl networks (Tugarinov et al. 2005; Sounier et al. 2007). Automated programs such as MAP-XS or FLAMEnGO utilize a pre-existing crystal structure of the protein to facilitate peak assignment based on predicted chemical shifts or an expected methyl–methyl NOESY cross peak network (Xu et al. 2009; Xu and Matthews 2013b; Chao et al. 2014). Although these programs are able to produce accurate assignments, the ambiguity present in the multidimensional NOESY datasets acquired with either ILV or ILVA labeled protein requires the programs to be supplemented with additional experimental information, including PRE data, data from through bond experiments or assignments derived from point mutations, in order to provide a more complete resonance assignment or to assign resonances with a higher degree of confidence. To address these limitations, the work described in this manuscript demonstrates the simultaneous, selective incorporation of isotope labels at all methyl positions (except the Ile Cγ position) of the six methyl containing amino acids (MILVAT). By extending the labeling strategy to include all six amino acids, we were able to use a single 4D ^{13}C-resolved HMQC–NOESY–HMQC experiment recorded on a single NMR sample to unambiguously assign all methyl groups in and around the enzymatic pocket of CoaD.

Materials and methods

Growth and expression of CoaD

Escherichia coli BL21(DE3) chemically competent cells (NEB product # C2530H) were transformed with 67 ng of pTrcHis2 B plasmid (Thermo Fisher Scientific) encoding the full length *E. coli* phosphopantetheine adenylyltransferase (PPAT) CoaD (Uniprot ID: P0A6I6) enzyme followed by an uncleavable, C-terminal His_6 tag. 5 mL nutrient rich LB medium containing carbenicillin (100 μg/mL) was inoculated with a single colony and grown until an optical density at 600 nm (OD_{600}) of 0.6 was achieved. At this point, 1 mL of LB culture was used to inoculate 50 mL of fully protonated M9 minimal medium, which was grown overnight at 37 °C with agitation. In all M9 growths, 1 g/L of ^{15}N-amonium chloride and 4 g/L of protonated ^{12}C-glucose were used as the principal nitrogen and carbon sources, respectively. *E. coli* cells were acclimated to growth in D_2O by growing the cells in 50 mL of M9 medium containing increasing concentrations of D_2O (30, 70 and 100 %) over a period of three days. On each occasion, 5 mL of the previous starter culture was used to inoculate the new growth which was left to grow at 37 °C overnight with agitation.

The 100 % D_2O overnight starter culture was used to inoculate 1 L of 100 % D_2O M9 medium to a starting OD_{600} of 0.1. *E. coli* cells were grown at 37 °C with agitation until an OD_{600} of 0.7 was achieved. At this stage, the temperature of the incubator was reduced to 18 °C and the desired combination of selectively labeled precursors and amino acids (Table 1), resuspended in 100 % D_2O, was added to the growth medium and left for 1 h. Following the 1 h incubation period, expression of CoaD was induced using 250 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) and cells were left growing at 18 °C for 18 h. Cells were harvested by centrifugation, washed with 1 × phosphate buffered saline (PBS) buffer and stored at −20 °C.

Purification of CoaD

Escherichia coli cells containing the overexpressed CoaD were resuspended in binding buffer (10 mL/g of cell pellet; 20 mM Tris pH 8.0, 500 mM sodium chloride, 1 mM tris(2-carboxyethyl)phosphine (TCEP), 20 mM imidazole), containing Roche Protease Inhibitor without ethylenediamine-tetraacetic acid (EDTA) and sonicated. Cell lysate was centrifuged at 40,000*g* for 30 min to remove insoluble cell debris, and the soluble lysate was passed through a 0.2 μm filter before being mixed with 5 mL Ni Sepharose 6 Fast Flow resin (GE Healthcare), which had been equilibrated in binding buffer. Soluble cell lysate and Ni resin

Table 1 Quantity of precursor or amino acid used to selectively label amino acids

Amino acid labeled	Precursor or amino acid used	Quantity (mg/L)	Comment and references
Methionine (Cε)	L-Methionine-(methyl-^{13}C)[a]	250	Fischer et al. (2007), Gelis et al. (2007)
Isoleucine (Cδ^1)	2-Ketobutyric acid-4-^{13}C,3,3-D$_2$[b]	70	Gardner and Kay (1997)
Leucine (Cδ^1, Cδ^2)	2-Keto-3-methyl-^{13}C-butyric-4-^{13}C,3-D acid[b]	120	Labeling of leucine Cδ atoms only can be achieved by supplementing the M9 medium with 20 % Bioexpress[a] (Goto et al. 1999; Tzeng et al. 2012)
Valine (Cγ^1, Cγ^2)	2-Keto-3-methyl-^{13}C-butyric-4-^{13}C,3-D acid[b]	120	Labeling of valine and leucine occurs simultaneously using the same precursor (Goto et al. 1999)
Alanine (Cβ)	L-Alanine (3-^{13}C,2-D)[a]	100	Ayala et al. (2009)
Threonine (Cγ)	L-Threonine (4-^{13}C;2,3-D$_2$)[a]	50	Isoleucine Cδ^1 atoms need to be labeled at the same time as threonine Cγ (Velyvis et al. 2012)
	Glycine (D$_5$)[a]	100	

Selectively labeled amino acids were purchased from [a] Cambridge Isotope Laboratory (product numbers: CLM-206, CDLM-9307, DLM-280, CGM-1000-D-S and CDLM-8649) and precursors for labeling isoleucine, leucine and valine residues were purchased from [b] Sigma (product numbers: 589063 and 589276)

were left on a mixing platform at room temperature for 1 h, before the resin was loaded into a disposable chromatography column (BioRad). Using gravity flow, the Ni resin was washed with 20 mL of wash buffer (20 mM Tris pH 8.0, 500 mM sodium chloride, 1 mM TCEP, 116 mM imidazole) and the CoaD was eluted from the column using 10 mL elution buffer (20 mM Tris pH 8.0, 500 mM sodium chloride, 1 mM TCEP, 500 mM imidazole).

To remove co-purified coenzyme A (CoA) from the CoaD active site, purified protein was concentrated to a volume of 500 µL in a 10 kDa Amicon Ultra centrifugal filter unit (Millipore) and resuspended to 10 mL using dialysis buffer (20 mM sodium citrate pH 5.0, 1 mM TCEP). This process was repeated three times. The protein solution was then placed into a 3 mL 10 kDa Slide-A-Lyzer dialysis cassette (Thermo Fisher Scientific) and dialyzed against 2 L of dialysis buffer for 48 h, with the buffer being exchanged every 24 h. The extent of CoA present in the sample was assessed by measuring the ratio of absorbance at 260 and 280 nm under native and thermal denaturing conditions (Geerlof et al. 1999).

The CoA free CoaD was exchanged into NMR buffer [20 mM sodium phosphate pH 7.0, 1 mM deuterated dithiothreitol (DTT)] using a PD-10 column (GE Healthcare) which had been pre-equilibrated in NMR buffer. Samples were concentrated to a final monomeric concentration of 350 µM, flash frozen in liquid nitrogen and stored at −80 °C until needed. Samples used for the 4D methyl–methyl ^{13}C-resolved HMQC–NOESY–HMQC experiment were lyophilized overnight and resuspended in an equal volume of 100 % D_2O (CIL).

NMR experiments

All NMR experiments were conducted with 160 µL of sample in a 3 mm NMR tube and were recorded at 308 K on a Bruker AVANCE III 600 MHz spectrometer equipped with a 5 mm CP-QCI z-gradient probe. Due to only labeling the methyl groups and along with the highly deuterated nature of our samples, we were able to utilize the increase in sensitivity obtained from measuring 2D [^{13}C, ^{1}H]-methyl TROSY experiments (Tugarinov et al. 2003). 2D [^{13}C, ^{1}H]-HMQC SOFAST (Schanda et al. 2005) experiments were acquired with an excitation flip angle of 100° and a proton band width of 3 ppm, centered at 1 ppm. Experiments were recorded using either uniform or 50 % non-uniform sampling with a recycle delay of 200 ms collecting a total of 1024 and 256 points in the direct and indirect dimensions, respectively. Spectra were processed using a QSINE window function, with 1.0 and 0.3 Hz line broadening applied in the direct and indirect dimensions, respectively. The 4D methyl–methyl ^{13}C-resolved HMQC–NOESY–HMQC experiment was recorded using 25 % non-uniform sampling (Tugarinov et al. 2005) with a mixing time of 120 ms. A total of 32,488 points were collected across all three indirect dimensions and 1024 points collected in the direct proton dimension. All non-uniformly sampled experiments were acquired with Poisson Gap sampling schedules (Hyberts et al. 2010), and were processed in Topspin 3.2 using either the hmsIST algorithm (Hyberts et al. 2012) for 2D-[^{13}C, ^{1}H]-SOFAST HMQC experiments or Multi-Dimensional Decomposition (MDD, Bruker) for the 4D ^{13}C-resolved HMQC–NOESY–HMQC experiment.

Calculation of equilibrium dissociation constants

K_d values for ligand binding were calculated by titrating between 12.5 μM and 5 mM ligand into a 50 μM sample of MILVAT labeled CoaD and measuring 2D [^{13}C, ^{1}H]-SOFAST HMQC experiments. The weighted change in chemical shift (Δ) of peaks, when compared to the apo spectrum of CoaD, was calculated using a ^{13}C scaling factor of 0.25 and the following formula was used to fit the data to calculate K_d values:

$$\Delta = \Delta_{max}\frac{\left([L]_T + [P]_T + K_d - \left\{\left([L]_T + [P]_T + K_d\right)^2 - 4[L]_T[P]_T\right\}^{\frac{1}{2}}\right)}{2[P]_T}$$

where Δ corresponds to the observed weighted change in chemical shift at a given total ligand concentration $[L]_T$, Δ_{max} is the change in chemical shift at saturation and $[P]_T$ is the total protein concentration.

Computation of methyl–methyl distances

Average distances between methyl groups of different amino acids of interest were computed using a Perl script. The script requires a protein data bank (PDB) file of the protein structure as input, as well as a list of the amino acids and the atoms of interest to be used in the analysis. The Perl script then determines carbon–carbon pairwise distances between all atoms of interest. Pairs of atoms separated by a distance above a user defined cut-off value are not considered. In the case of multimeric assemblies, distances for the same atom pairs in different subunits are averaged to account for small structural variations.

Results and discussion

The protein of interest used in this study is the *E. coli* phosphopantetheine adenylyltransferase CoaD, which natively exists as a 115 kDa homohexamer (Fig. 1). CoaD catalyzes the magnesium dependent transfer of an adenylyl group from ATP to 4′-phosphopantetheine to generate dephospho-CoA, which is an intermediate in the biosynthesis of coenzyme A (CoA) (Geerlof et al. 1999). CoA is an essential enzyme cofactor in all living organisms and plays a central role in reactions involved with cellular metabolism as well as fatty acid, polyketide and non-ribosomal peptide biosynthesis (Leonardi et al. 2005; Spry et al. 2008). For this reason and due to the limited conservation between the bacterial and human forms of the enzyme, CoaD is also of interest as a potential antibacterial drug target (Aghajanian and Worrall 2002; Daugherty et al. 2002; Miller et al. 2007). Analysis of the protein amino acid sequence shows that methyl containing amino acids constitute 44 % of the polypeptide sequence and are distributed evenly throughout the structure (Fig. 1). The enzyme has been both biochemically and structurally characterized exhaustively, and as a result a number of structures, both in the absence and presence of inhibitors as well as natural products are already available in the PDB (Izard and Geerlof 1999; Izard 2002, 2003; Miller et al. 2007).

Production of MILVAT labeled protein

The goal of the research presented here was to produce a protocol to simultaneously label all six methyl containing amino acids. In order to see if this could be achieved, CoaD samples where the six methyl containing amino acids were labeled individually (M, I, A) or in their respective pairs (LV, IT) were produced as a benchmark for comparison to a MILVAT labeled sample. A sample of LV protein was also produced in the presence of 20 % BioExpress to suppress the incorporation of label at leucine Cδ positions (Tzeng et al. 2012). [^{13}C, ^{1}H]-SOFAST-HMQC experiments sampling the methyl region were acquired using 50 μM samples of each selectively labeled protein. The resulting individual spectra are shown in Supplementary Fig. 1, and a composite HMQC spectrum containing a superposition of the individual spectra, colored according to amino acid type, is shown in Fig. 2a. All 103 expected peaks are present in the spectra and only a minimal amount of overlap is observed between the leucine δ^1 and δ^2 and valine γ^1 and γ^2 resonances.

A MILVAT labeled protein sample was produced by resuspending all six precursors and labeled amino acids (Table 1) in D_2O and adding this solution to the *E. coli* growth 1 h prior to induction. A [^{13}C, ^{1}H]-SOFAST-HMQC experiment acquired with 50 μM MILVAT labeled protein (Fig. 2b) shows that simultaneous labeling of all six amino acid methyl groups achieves an identical HMQC spectrum to the composite spectrum produced from the individually labeled protein samples (Fig. 2a). In addition, a [^{13}C, ^{1}H]-HMQC spectrum covering the full aliphatic and aromatic regions of the spectrum was acquired with folding in the indirect dimension (Supplementary Fig. 2). Even when analyzed at the level of the noise, the only significant peaks present in the spectrum are those of the resonances specifically labeled using MILVAT labeling. A single threonine resonance (^{1}H 0.96 ppm, ^{13}C 21.31 ppm), which overlaps with two leucine δ resonances, is not resolved in the MILVAT spectrum, however the presence of a threonine peak at this position was confirmed using a 4D HMQC–NOESY–HMQC experiment.

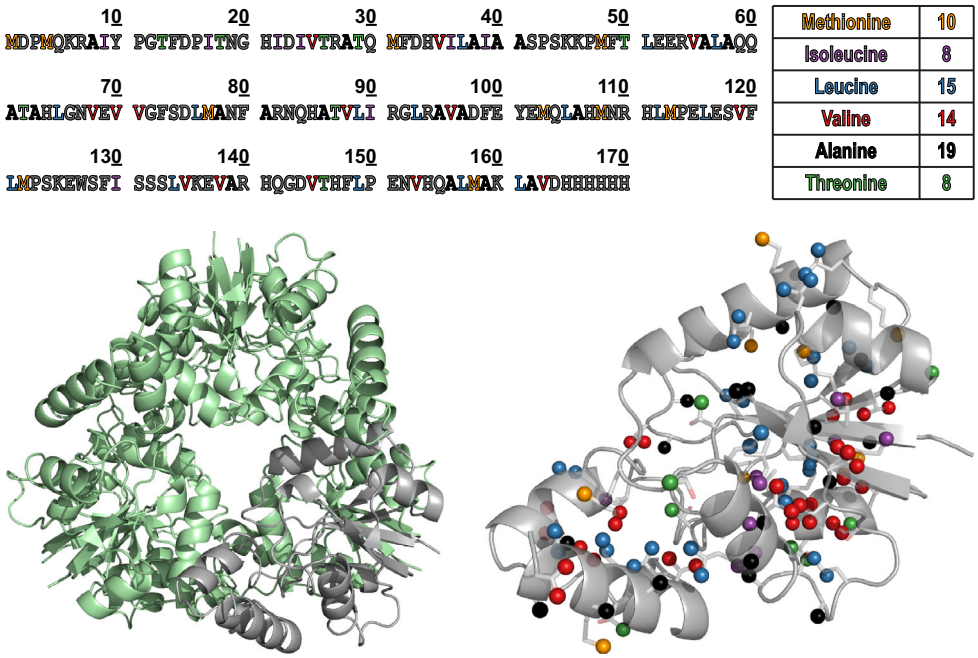

Fig. 1 Analysis of the amino acid sequence and structural elements of CoaD. *Top panel* Amino acid sequence of *E. coli* CoaD. A summary of the number of methyl containing amino acids present in the sequence is provided in the table to the right of the sequence. Amino acids are colored using the color scheme used in the table (*Ile purple, Met orange, Ala black, Thr green, Val red and Leu blue*). *Bottom panel* Structure of the *E. coli* CoaD (PDB entry 5JBN) homohexamer shown as a cartoon representation. One CoaD protomer is shown in *light grey*, with the additional five other subunits are colored *light green*. One CoaD subunit showing the location of all NMR active probes present in a MILVAT labeled protein sample is depicted on the *right*. *Spheres* representing the NMR active nuclei are colored according to the color scheme detailed above

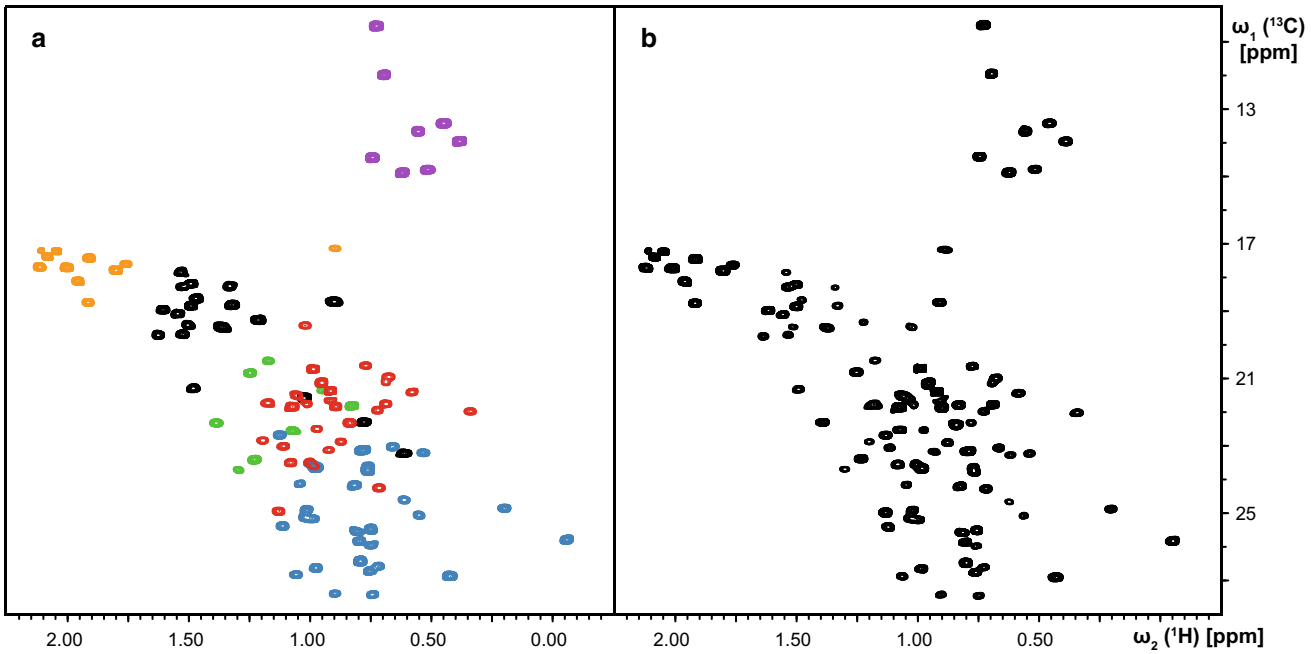

Fig. 2 2D [^{13}C, ^{1}H]-HMQC spectra of *E. coli* CoaD. **a** A composite HMQC spectrum was prepared by superimposing all 2D [^{13}C, ^{1}H]-SOFAST HMQC spectra acquired with individually labeled CoaD samples (see main text and Supplementary Fig. 1). Peaks have been colored according to the amino acid color scheme outlined in Fig. 1. **b** A 2D [^{13}C, ^{1}H]-SOFAST HMQC spectrum of MILVAT labeled CoaD is shown. Protein concentrations and buffer conditions were the same for all samples, which were acquired at 308 K at 600 MHz ^{1}H frequency

Analysis of scrambling into unwanted methyl positions

Previous reports published by Ayala et al. and Kerfah et al. concluded that scrambling into multiple amino acid side chains, including isoleucine at the $C\gamma^2$ position, was observed when supplementing the medium with alanine, selectively labeled at the Cβ position (Ayala et al. 2009; Kerfah et al. 2015b). Figure 3a shows the HMQC spectrum of a MILVAT labeled sample, along with a magnified area corresponding to the region of the spectrum where we would expect to see isoleucine $C\gamma^2$ resonances. Even when plotting the spectrum at the level of the noise, no unexpected peaks are present in the spectrum, indicating that there is no scrambling into the isoleucine $C\gamma^2$ position in the MILVAT labeled protein sample. Even though inclusion of 100 mg/L of selectively labeled alanine during expression is sufficient to label all alanine Cβ resonances (see below), the lack of scrambling could have been attributed to the relatively low concentration of labeled alanine included during protein expression (Table 1). To address this, we performed similar analyses with protein samples expressed with higher concentrations of labeled alanine (up to 700 mg/L), and again no undesired incorporation into the isoleucine $C\gamma^2$ position was observed. Scrambling has also been reported to occur from alanine into the valine and leucine Cγ and Cδ positions, respectively (Ayala et al. 2009; Godoy-Ruiz et al. 2010; Kerfah et al. 2015b); however, we are not concerned about any potential scrambling into these positions as these atoms are intentionally labeled in this protocol.

Probing the extent of isotope incorporation

In addition, Ayala et al. also observed a correlation between the concentration of labeled alanine added prior to induction and the incorporation of selective label at the alanine Cβ position (Ayala et al. 2009; Kerfah et al. 2015b). The concentration dependence of alanine required to achieve complete incorporation in our system was addressed using ILVA growths where different concentrations of alanine (100, 200 and 700 mg/L) were supplemented into the medium 1 h prior to induction. All samples were purified in the same way and 2D [^{13}C, ^{1}H]-SOFAST HMQC experiments were acquired on protein solutions of the same concentration using identical acquisition parameters. The peaks corresponding to alanine residues were picked in all three spectra and their absolute intensities were determined (Fig. 3b). There is no difference observed between the intensities of alanine residues in samples which were supplemented with different concentrations of alanine prior to induction, indicating that the same degree of alanine incorporation is achieved in all three samples. With this in mind, we decided to supplement our minimal medium with 100 mg/L of selectively labeled alanine for all future experiments. When comparing the different protocols, differences are apparent in the carbon sources used for bacterial growth and protein expression. Ayala et al. use protonated glycerol as the sole carbon source, while Kerfah et al. use deuterated ^{12}C glucose (Ayala et al. 2009; Kerfah et al. 2015b). Our protocol uses protonated ^{12}C glucose as the predominant carbon source (additional carbon is supplemented into the medium in the form of selectively labeled precursors and amino acids) and although this may slightly decrease the total extent of deuteration of the protein, it does not appear to significantly affect the quality of the spectrum produced. Although this is a major difference between the three protocols, there are also other subtle variations in the expression protocols including concentrations of IPTG used along with the applied temperatures and lengths of time used for protein induction which could also contribute to the differences observed.

To quantitatively assess overall isotope incorporation, we performed mass spectrometry analysis on the MILVAT labeled CoaD protein. The expected mass is 20,239.8 Da, assuming that all amide bound protons have exchanged with H_2O and that all methyl containing amino acids are fully labeled. The experimentally determined molecular weight peaks at 20,143.7 Da (Supplementary Fig. 3), which equates to a difference in molecular weight of 96 Da. This corresponds to approximately one proton being incorporated into each amino acid that is not selectively labeled on average. An asymmetric mass distribution of approximately 140 Da is observed around the average mass, which presumably arises from different degrees of amide back exchange along with incorporation of protons from water present in D_2O and the protonated glucose. Our experimental data indicates that isotope incorporation of greater than 99 % can be achieved using the outlined protocol, however, if higher levels are required the use of deuterated ^{12}C-glucose could be considered.

To analyze if preferential incorporation of supplemented amino acids or precursors into one amino acid over another was obtained, all peaks present in the [^{13}C, ^{1}H]-HMQC spectrum of the MILVAT labeled protein (Fig. 2b) were picked and their intensities were determined. Peaks were sorted according to their intensity and a profile for all peaks was generated (Fig. 3c; Pedrini et al. 2013). With exception of the first 15 peaks, which have a high intensity as a result of being surface exposed or present in a mobile area of the protein, all CoaD methyl peaks have a fairly even intensity, and do not abnormally cluster according to amino acid type in the intensity profile. This indicates that we see an equal incorporation of isotopes into all six amino acid types.

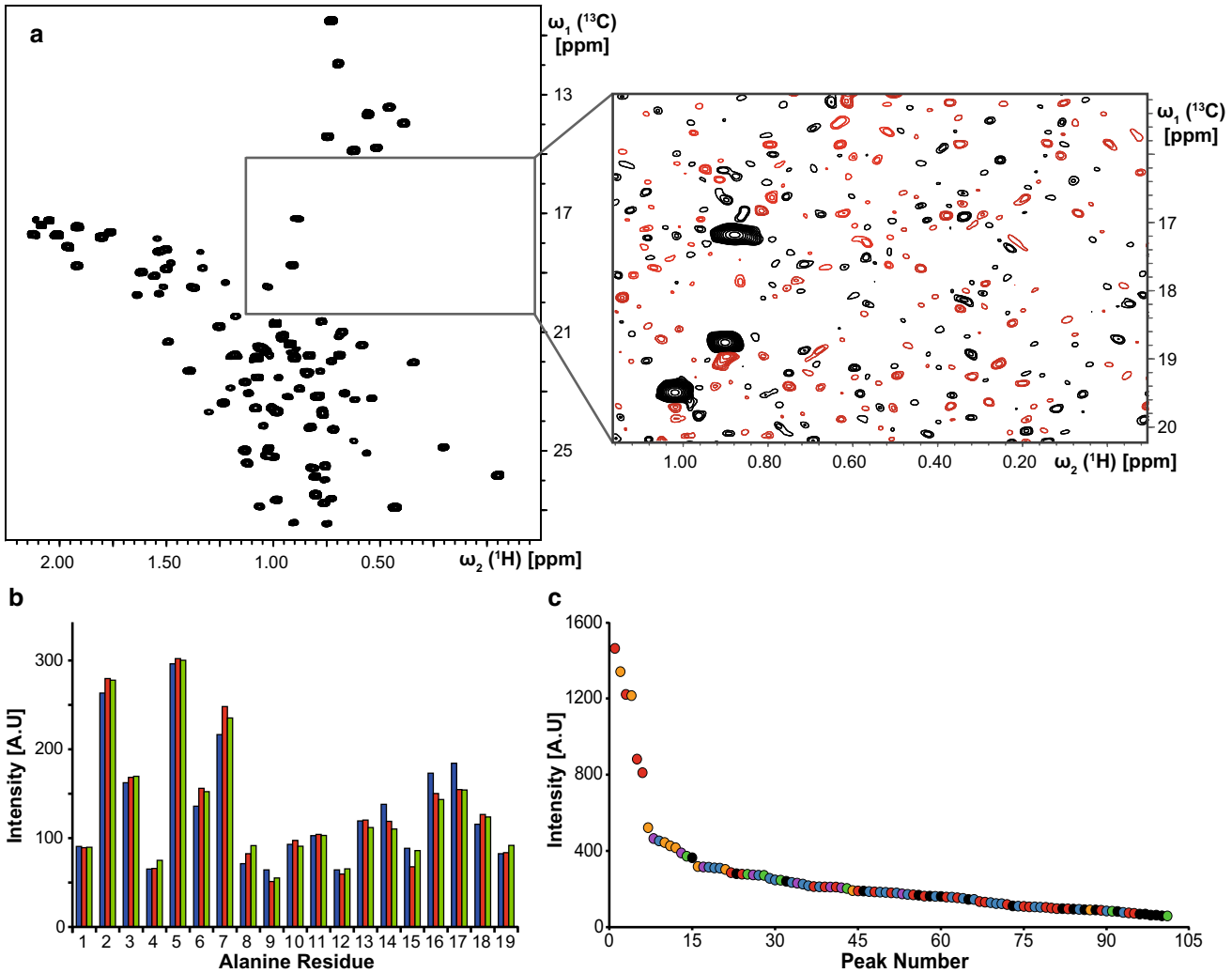

Fig. 3 Characterization of amino acid scrambling and incorporation efficiencies. **a** 2D [^{13}C, ^{1}H]-SOFAST HMQC spectrum of MILVAT labeled CoaD, highlighting the region where Ile Cγ^2 resonances would appear in the spectrum. Lack of signal in this area demonstrates that no scrambling was observed into this position. **b** Comparison of peak intensities of alanine residues present in spectra acquired on different ILVA samples expressed with different L-alanine (3-^{13}C, 2-D) concentrations. *Blue*, *red* and *green* bars correspond to peaks present in protein samples expressed when the medium was supplemented with 100, 200 and 700 mg/L of selectively labeled alanine one hour prior to induction, respectively. The equal height of the *bars* indicates that for the range of concentrations tested, the incorporation of alanine into the protein is not influenced by the amount of amino acid used during the expression. **c** Intensity profile of all peaks present in the 2D [^{13}C, ^{1}H]-SOFAST HMQC spectrum acquired on MILVAT labeled CoaD. All peaks were picked and maximum intensities were determined using Cara (Keller 2004). Data points on the graph are colored according to the color scheme detailed above. Apart from the first 15 peaks which have a high intensity due to being surface exposed or being in an area of the protein with high mobility, all peaks have a fairly even intensity and do not cluster according to amino acid type, indicating an equal incorporation of isotopes into all six labeled amino acids

When compared to the individually labeled protein samples, a small decrease in methyl signal intensity was observed in the case of CoaD upon labeling all six amino acids. This effect could be attributed to line broadening due to enhanced relaxation caused by the increased number of NMR active probes in the sample (e.g. illustrated in Kerfah et al. 2015c). To assess this, the linewidths of all sufficiently isolated peaks were determined in the extracted 1D rows of the HMQC spectra. When all six amino acids were labeled individually, the average linewidth of all peaks analyzed was 16.0 Hz. As expected, upon increasing the number of labels present in the sample, line broadening is observed for all amino acid types (Table 2). When compared to the single labeled samples (M, I, A), peaks corresponding to the same residues in the MILVAT labeled protein sample have an increased linewidth of approximately 1.8 Hz (Table 2). For samples where amino acids are labeled as pairs (LV, IT), the increase in linewidth

Table 2 Line shape analysis of peaks present in individually and combinatorially labeled CoaD samples

Amino acid	Labeling scheme	# Peaks analyzed	Average absolute linewidth (Hz)	Average change in linewidth with different labeling schemes (Hz)		
				ILV	ILVA	MILVAT
M	M	7	15.5			1.6
I	I	8	14.4	0.8	1.3	1.7
L	LV	14	17.2	−0.1	0.2	0.7
V	LV	12	16.5	−0.2	0.1	1.0
A	A	12	16.3		1.1	2.0
T	IT	7	16.1			0.7

observed in the MILVAT labeled protein is approximately 0.7 Hz. A respective increase in linewidth between 0.8–1.0 and 0.4–0.8 Hz is observed for peaks when comparing MILVAT labeling to the standard ILV and ILVA labeling schemes (Table 2). Although these numbers are expected to vary between different protein systems, the data presented here suggests that the benefits obtained from having all six methyl groups labeled simultaneously can outweigh the increase in spin relaxation, especially when compared to that already observed with an ILVA labeling scheme.

Using the described protocol, to date we have successfully labeled and analyzed multiple homo-multimeric proteins with the size of the molecular assemblies ranging between 45 and 150 kDa. For all proteins analyzed, there was no scrambling into unwanted positions and a general increase in signal linewidth was observed as detailed above. However, in some of the systems studied, a larger decrease in signal intensity was observed for a subset of threonine signals when compared to the decrease observed for other peaks in the spectrum. This may be a combined result of the increased proton density along with the dynamics and conformational exchange occurring in the proteins analyzed (Velyvis et al. 2012).

Assignment of methyl groups using a single 4D ^{13}C-resolved HMQC–NOESY–HMQC experiment

When selective labeling of methyl groups in proteins is required, the standard labeling schemes used are either ILV or ILVA labeling. A comparison of MILVAT labeling to these standard labeling protocols indicates that the increase in probe coverage obtained from selectively labeling all six methyl containing amino acids would facilitate assignment of all methyl groups using a single 4D ^{13}C-resolved HMQC–NOESY–HMQC experiment. Previously, the strategy of consecutively assigning neighboring methyl groups through NOE connectivities ('methyl walk') might have been limited since many distances between labeled probes exceed the detection limit of a NOESY experiment. In the MILVAT labeled protein sample, NMR active nuclei are now much more likely to be present in these gaps, thus allowing the user to 'walk' from methyl to methyl and assign the peaks based on the known structure of the protein. To demonstrate the feasibility of an efficient methyl walk using MILVAT-labeled protein, a 4D ^{13}C-resolved HMQC–NOESY–HMQC experiment, which had been resolution optimized so only the methyl region in the indirect proton dimension was sampled, was implemented. Using the crystal structure of apo CoaD (PDB entry 5JBN), the intra- and inter-molecular distances between the carbon atoms of methyl groups in the protein monomer and any symmetry pairs in the oligomeric unit were calculated. An upper distance limit of 7.5 Å was imposed while conducting the analysis to simulate the expected maximum nuclear Overhauser effect (NOE) observable distance, as well as to accommodate for the additional distance of the attached proton atoms, which is not accounted for in the analysis.

The 4D methyl–methyl ^{13}C-resolved HMQC–NOESY–HMQC experiment was analyzed by evaluating cross peaks present in the (F3, F4) HMQC plane with the F2 and F1 dimensions set to the frequencies of the methyl resonance being assessed (Fig. 4). Using the beforehand determined amino acid type assignments, we were able to go through the spectrum and generate a network of interactions observed by each methyl group in the sample. By comparing this list of observed methyl NOEs to the output of expected methyl NOEs, we were able to find multiple patterns which are unique to particular methyl groups. Such patterns then served as starting points from which we could conduct a series of methyl walks to assign the surrounding methyl groups (Fig. 4). For the purposes of the work being conducted, the complete assignment of the protein was not required, however using a single MILVAT labeled sample, along with the information obtained from the individually labeled protein preparations, we were able to confidently assign 60 % of all methyl groups in CoaD, which included the complete enzymatic pocket as well as residues immediately surrounding that area, within one day. Interestingly, all six different methyl containing amino acids are found in

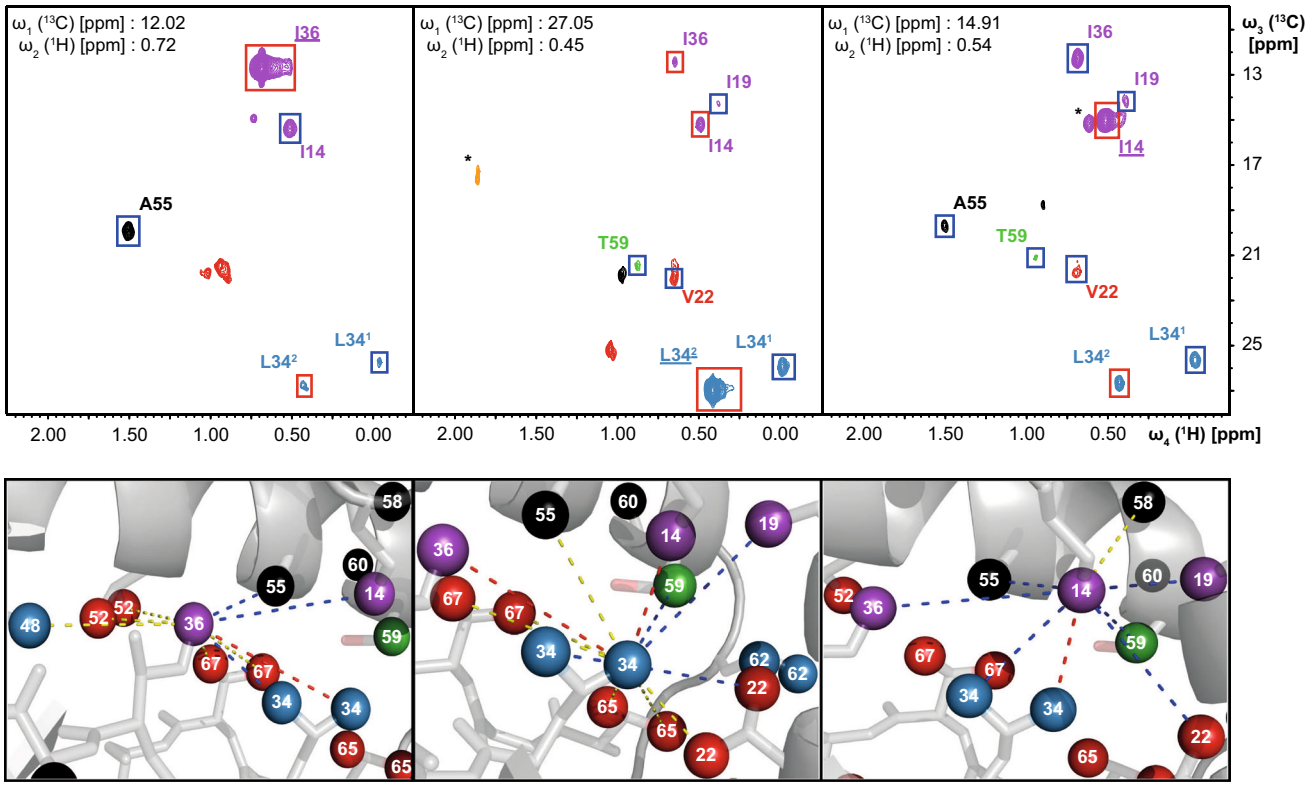

Fig. 4 An example of a 'methyl walk' used to assign selectively labeled methyl groups in CoaD. *Upper panel* contains 2D [^{13}C, ^{1}H]-planes of the 4D methyl–methyl ^{13}C-resolved HMQC–NOESY–HMQC experiment recorded with MILVAT labeled CoaD. Peaks are colored according to their amino acid type using the color scheme detailed above. *Red boxes* are drawn around resonance peaks used in the methyl walk shown, whereas *blue boxes* highlight redundant resonance peaks that appear in multiple planes which could be used to generate unambiguous assignments. Assignments that are underlined correspond to peaks that have the ^{1}H, ^{13}C chemical shifts detailed at the top left of each spectrum for which NOEs are being observed to. Peaks marked with an *asterisk* correspond to those that have their maximum intensity in a different plane and for which residual intensity is observed in the plane shown. Peaks that are not labeled correspond to resonances that could not be assigned using the information present in the methyl walk shown. *Lower panel* contains cartoon representations of the structure of apo CoaD (PDB entry 5JBN) centered on methyl groups that have underlined assignments. All methyl groups are represented by *spheres* and are colored using the scheme detailed above, and methyl groups within a 7.5 Å radius of the centered atom are connected by *colored lines*. *Red* and *blue lines* correspond to the *boxes in the upper part* of the figure, while *yellow lines* correspond to methyl groups within 7.5 Å that could not be unambiguously assigned in the methyl walk shown

the active site (Supplementary Fig. 4) and the simultaneous labeling of all amino acids greatly facilitated the assignment due to the absence of gaps in the methyl–methyl network.

Interaction studies with (R)-pantetheine

In addition to facilitating resonance assignment, the increased probe density in a MILVAT labeled protein also enables the study of structure–function relationships with greater detail and accuracy. For example, the binding site of a ligand could be more comprehensively captured in a single mapping experiment. To demonstrate this, we assessed the binding of the substrate analogue (R)-pantetheine by monitoring chemical shift perturbations of protein methyl groups. A crystal structure of *E. coli* CoaD in complex with 4′-phosphopantetheine [PDB entry 1QJC, (Izard 2002)] shows that the binding site covers one side of the complete enzymatic pocket which is partially comprised of residues from the neighboring subunit in the hexameric holoenzyme (Fig. 5a). Since the reported K_d for phosphopantetheine is 270 nM (Miller et al. 2007), we chose the analogue (R)-pantetheine for our binding study in order to form a weaker interaction which would be in fast exchange. A 50 μM sample of MILVAT labeled CoaD was measured with increasing concentrations of (R)-pantetheine up to a final concentration of 5 mM. Addition of (R)-pantetheine to the sample induced chemical shift perturbations in the spectrum, indicating that the substrate analogue interacted with CoaD (Fig. 5b) in the fast exchange regime. The changes in chemical shift of multiple peaks in the spectrum were used to determine a K_d of 1.6 ± 0.3 mM. Analysis of the assignments of peaks in the spectrum which exhibited the largest changes in chemical

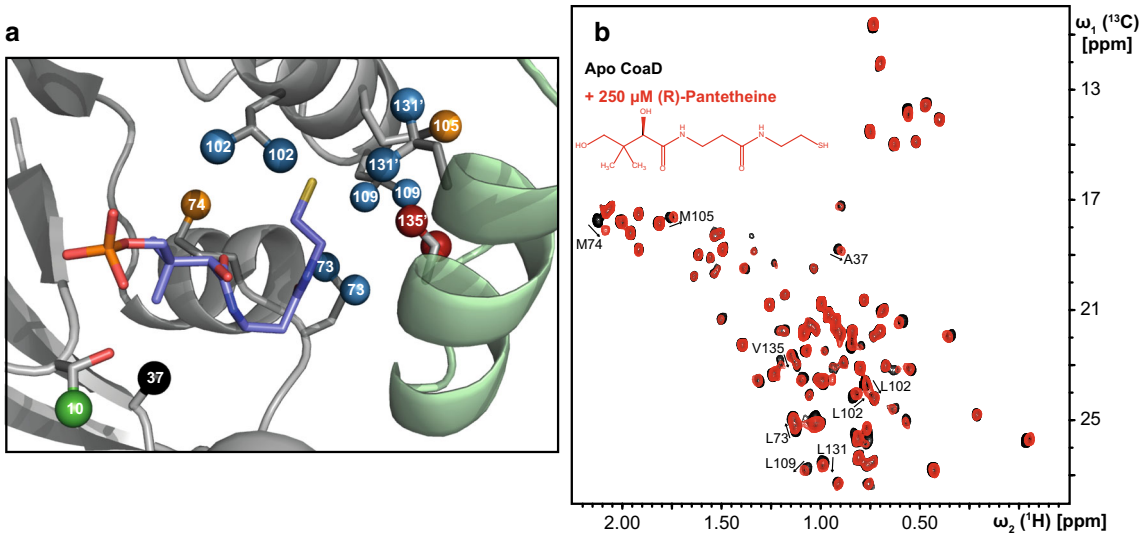

Fig. 5 Analysis of the binding of (R)-pantetheine to *E. coli* CoaD. **a** Cartoon representation of the structure of CoaD in complex with 4′-phosphopantetheine (PDB entry 1QJC). The two different subunits which make up the active site have been colored in *light grey* and *light green*, respectively. All methyl groups present in both subunits that are within a 6.5 Å radius of the 4′-phosphopantetheine are represented as *spheres* and colored according to the color scheme detailed above. **b** Superposition of 2D [^{13}C, ^{1}H]-SOFAST HMQC spectra of MILVAT labeled CoaD recorded in the absence (*black*) and presence of 250 µM (R)-pantetheine (*red*). Peaks which undergo chemical shift perturbations upon the addition of (R)-pantetheine have been labeled and correspond to methyl groups in the phosphopantetheine binding site

shift showed that methionine, leucine, valine and alanine residues are affected and that those peaks corresponded to methyl groups which are directly surrounding the ligand. Additional smaller perturbations are observed in second shell and more remote residues upon addition of a ligand which is consistent with the observations made by others that CoaD is an allosteric enzyme (Izard 2002). The increased number and type of affected peaks observed by employing the new labeling scheme will allow characterization of these processes in more detail, e.g. by magnetization relaxation based dynamics studies.

Conclusions

We have developed a protocol for the selective, simultaneous labeling of all methyl groups (except Ile Cγ) in the six methyl containing amino acids and have demonstrated how such a sample can be used to facilitate the assignment of methyl groups in the enzyme CoaD. The simultaneous labeling of all six amino acids does not adversely affect the spectral quality of CoaD and the addition of selectively labeled threonine, alanine and methionine resonances hardly increases the amount of peak overlap due to the unique places in the spectrum where these resonances appear. There are significant benefits of observing all six amino acids simultaneously in the spectrum, especially when using NOESY datasets to assign peaks, since the unique chemical shift fingerprints of peaks obtained from having methionine, alanine and threonine labeled help when trying to identify unique starting points or when trying to unambiguously assign peaks.

There are currently a number of different computational programs that conduct the assignment of methyl groups using the information from multi-dimensional NOESY datasets, however, these programs generally need to be supplemented with additional experimental data in order to increase the quantity of or confidence in the unambiguous assignments made (Chao et al. 2014). It is our belief that the use of both 4D spectroscopy and MILVAT labeling would greatly increase the performance of programs which work to produce a complete methyl assignment based on identifying networks of methyl–methyl NOEs in multidimensional NOESY datasets, without having to supplement the analysis with additional experimental data.

Isolated methyl groups that are not present in a methyl–methyl network cannot be assigned directly using this process, however if these groups are required for the analysis being conducted, often they can be assigned through the process of elimination. In most cases where structures are not being calculated, a complete assignment of all methyl groups is not needed and research groups are only interested in the subset of peaks from residues within a protein active site or binding interface which are involved in the reaction or interaction being studied. The use of MILVAT labeling offers multiple unique starting points in

the 4D NOESY dataset from which methyl walks can be conducted, facilitating the assignment of only the regions of interest. When compared to the standard ILV or ILVA labeled samples, the percentage of methyl groups labeled in a MILVAT sample increases on average to 35 % from 18 and 27 %, respectively (Miller et al. 1987; Conte et al. 1999). This increased number of NMR active probes in the sample significantly increases the redundancy observed in the NOESY dataset, thus providing increased confidence in the assignments made. In addition to the total increase in the number of NMR active probes present in the sample, a 2.3 or 1.4 fold increase in the number of probes available on the protein surface is observed in a MILVAT labeled sample, when compared to the standard ILV or ILVA labeling scheme, respectively (Miller et al. 1987). Extending selective labeling to include methionine and threonine is of particular interest because of the unique flexible aliphatic thioether moiety and polar hydroxyl group in their respective side chains. These features often allow for productive interactions on the protein surface and in ligand binding sites, whereas isoleucine, leucine and valine side chains are generally found in the hydrophobic core.

The increased number of probes in a protein's active site enables a more detailed structural characterization of molecular interactions, which is of particular interest when studying protein small molecule complexes due to the highly localized binding event. Additional intermolecular NOEs measured to unambiguously assigned methyl groups are important to reduce the possible binding modes and ligand conformations in order to obtain models with reasonable confidence. As NOE-derived intermolecular distance information is preferably obtained from a single sample, MILVAT labeling is essential for any form of structural analysis using NOESY experiments. In addition to the benefits offered for such structural and interaction studies, the use of MILVAT labeling is also advantageous for groups interested in protein dynamics, conformational analysis, methyl detected relaxation dispersion and paramagnetic relaxation enhancement. Although for such experiments similar results could be achieved from studying multiple different samples, MILVAT labeling allows all information to be collected simultaneously and analyzed using a single sample, thus leading to both financial and time savings as well as to a simplification of analysis without having to compensate for sample variation.

Acknowledgments The authors would like to thank Dr. Clemens Anklin and Dr. Eric Johnson at the Bruker BioSpin Corporation for their assistance with the implementation of the 4D methyl–methyl ^{13}C-resolved HMQC–NOESY–HMQC experiment, Dr. Lili Xie and Kyoko Uehara for their help in designing the purification protocol for CoaD, and Dr. Dirksen Bussiere for his assistance with refinement of the CoaD structure. All images produced of the CoaD structure were done so using Pymol (Schrödinger, LLC 2015).

Compliance with ethical standards

Conflict of interest AP, AOF, FR, MM and AL are employees of Novartis.

References

Aghajanian S, Worrall DM (2002) Identification and characterization of the gene encoding the human phosphopantetheine adenylyltransferase and dephospho-CoA kinase bifunctional enzyme (CoA synthase). Biochem J 365:13–18. doi:10.1042/BJ20020569

Ayala I, Sounier R, Usé N et al (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. J Biomol NMR 43:111–119. doi:10.1007/s10858-008-9294-7

Chao F-A, Kim J, Xia Y et al (2014) FLAMEnGO 2.0: an enhanced fuzzy logic algorithm for structure-based assignment of methyl group resonances. J Magn Reson 245:17–23. doi:10.1016/j.jmr.2014.04.012

Conte LL, Chothia C, Janin J (1999) The atomic structure of protein–protein recognition sites. J Mol Biol 285:2177–2198. doi:10.1006/jmbi.1998.2439

Crublet E, Kerfah R, Mas G et al (2014) A cost-effective protocol for the parallel production of libraries of $^{13}CH_3$-specifically labeled mutants for NMR studies of high molecular weight proteins. In: Chen YW (ed) Structural genomics. Humana Press, Totowa, pp 229–244

Daugherty M, Polanuyer B, Farrell M et al (2002) Complete reconstitution of the human coenzyme A biosynthetic pathway via comparative genomics. J Biol Chem 277:21431–21439. doi:10.1074/jbc.M201708200

Fischer M, Kloiber K, Häusler 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. doi:10.1002/cbic.200600551

Gardner KH, Kay LE (1997) Production and incorporation of ^{15}N, ^{13}C, ^{2}H (^{1}H-δ1 methyl) isoleucine into proteins for multidimensional NMR studies. J Am Chem Soc 119:7599–7600. doi:10.1021/ja9706514

Geerlof A, Lewendon A, Shaw W (1999) Purification and characterization of phosphopantetheine adenylyltransferase form *Escherichia coli*. J Biol Chem 274:27105–27111. doi:10.1074/jbc.274.38.27105

Gelis I, Bonvin AMJJ, Keramisanou D et al (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell 131:756–769. doi:10.1016/j.cell.2007.09.039

Godoy-Ruiz R, Guo C, Tugarinov V (2010) Alanine methyl groups as NMR probes of molecular structure and dynamics in high-molecular-weight proteins. J Am Chem Soc 132:18340–18350. doi:10.1021/ja1083656

Goto NK, Kay LE (2000) New developments in isotope labeling strategies for protein solution NMR spectroscopy. Curr Opin Struct Biol 10:585–592. doi:10.1016/S0959-440X(00)00135-4

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-, ^{2}H-labeled proteins. J Biomol NMR 13:369–374. doi:10.1023/A:1008393201236

Gross JD, Gelev VM, Wagner G (2003) A sensitive and robust method for obtaining intermolecular NOEs between side chains in large protein complexes. J Biomol NMR 25:235–242. doi:10.1023/A:1022890112109

Hajduk PJ, Augeri DJ, Mack J et al (2000) NMR-based screening of proteins containing ^{13}C-labeled methyl groups. J Am Chem Soc 122:7898–7904. doi:10.1021/ja000350l

Hyberts SG, Takeuchi K, Wagner G (2010) Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. J Am Chem Soc 132:2145–2147. doi:10.1021/ja908004w

Hyberts SG, Milbradt AG, Wagner AB et al (2012) Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. J Biomol NMR 52:315–327. doi:10.1007/s10858-012-9611-z

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. doi:10.1021/ja0761784

Izard T (2002) The crystal structures of phosphopantetheine adenylyltransferase with bound substrates reveal the enzyme's catalytic mechanism. J Mol Biol 315:487–495. doi:10.1006/jmbi.2001.5272

Izard T (2003) A novel adenylate binding site confers phosphopantetheine adenylyltransferase interactions with coenzyme A. J Bacteriol 185:4074–4080. doi:10.1128/JB.185.14.4074-4080.2003

Izard T, Geerlof A (1999) The crystal structure of a novel bacterial adenylyltransferase reveals half of sites reactivity. EMBO J 18:2021–2030. doi:10.1093/emboj/18.8.2021

Keller R (2004) The computer aided resonance assignment tutorial. Cantina Verlag, Goldau

Kerfah R, Hamelin O, Boisbouvier J, Marion D (2015a) CH_3-specific NMR assignment of alanine, isoleucine, leucine and valine methyl groups in high molecular weight proteins using a single sample. J Biomol NMR 63:389–402. doi:10.1007/s10858-015-9998-4

Kerfah R, Plevin MJ, Pessey O et al (2015b) Scrambling free combinatorial labeling of alanine-β, isoleucine-δ^1, leucine-proS and valine-proS methyl groups for the detection of long range NOEs. J Biomol NMR 61:73–82. doi:10.1007/s10858-014-9887-2

Kerfah R, Plevin MJ, Sounier R et al (2015c) Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. Curr Opin Struct Biol 32:113–122. doi:10.1016/j.sbi.2015.03.009

Leonardi R, Zhang Y, Rock C, Jackowski S (2005) Coenzyme A: back in action. Prog Lipid Res 44:125–153. doi:10.1016/j.plipres.2005.04.001

Lichtenecker R, Ludwiczek ML, Schmid W, Konrat R (2004) Simplification of protein NOESY spectra using bioorganic precursor synthesis and NMR spectral editing. J Am Chem Soc 126:5348–5349. doi:10.1021/ja049679n

Miller S, Janin J, Lesk AM, Chothial C (1987) Interior and surface of monomeric proteins. J Mol Biol 196:641–656. doi:10.1016/0022-2836(87)90038-6

Miller JR, Ohren J, Sarver RW et al (2007) Phosphopantetheine adenylyltransferase from *Escherichia coli*: investigation of the kinetic mechanism and role in regulation of coenzyme A biosynthesis. J Bacteriol 189:8196–8205. doi:10.1128/JB.00732-07

Pedrini B, Serrano P, Mohanty B et al (2013) NMR-profiles of protein solutions. Biopolymers 99:825–831. doi:10.1002/bip.22348

Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci USA 94:12366–12371. doi:10.1073/pnas.94.23.12366

Religa TL, Kay LE (2010) Optimal methyl labeling for studies of supra-molecular systems. J Biomol NMR 47:163–169. doi:10.1007/s10858-010-9419-7

Religa TL, Sprangers R, Kay LE (2010) Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. Science 328:98–102. doi:10.1126/science.1184991

Saio T, Guan X, Rossi P et al (2014) Structural basis for protein antiaggregation activity of the trigger factor chaperone. Science 344:1250494. doi:10.1126/science.1250494

Schanda P, Kupče Ē, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. J Biomol NMR 33:199–211. doi:10.1007/s10858-005-4425-x

Schrödinger LLC (2015) The {PyMOL} molecular graphics system, Version ~1.8

Sounier R, Blanchard L, Wu Z, Boisbouvier J (2007) High-accuracy distance measurement between remote methyls in specifically protonated proteins. J Am Chem Soc 129:472–473. doi:10.1021/ja067260m

Sprangers R, Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature 445:618–622. doi:10.1038/nature05512

Spry C, Kirk K, Saliba KJ (2008) Coenzyme A biosynthesis: an antimicrobial drug target. FEMS Microbiol Rev 32:56–106. doi:10.1111/j.1574-6976.2007.00093.x

Stoffregen MC, Schwer MM, Renschler FA, Wiesner S (2012) Methionine scanning as an NMR tool for detecting and analyzing biomolecular interaction surfaces. Structure 20:573–581. doi:10.1016/j.str.2012.02.012

Tugarinov V, Kay LE (2003a) Side chain assignments of Ile δ^1 methyl groups in high molecular weight proteins: an application to a 46 ns tumbling molecule. J Am Chem Soc 125:5701–5706. doi:10.1021/ja021452+

Tugarinov V, Kay LE (2003b) 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. doi:10.1021/ja030345s

Tugarinov V, Hwang PM, Ollerenshaw JE, Kay LE (2003) Cross-correlated relaxation enhanced ^{1}H-^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. J Am Chem Soc 125:10420–10428. doi:10.1021/ja030153x

Tugarinov V, Hwang PM, Kay LE (2004) Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. Annu Rev Biochem 73:107–146. doi:10.1146/annurev.biochem.73.011303.074004

Tugarinov V, Kay LE, Ibraghimov I, Orekhov VY (2005) High-resolution four-dimensional ^{1}H-^{13}C NOE spectroscopy using methyl-TROSY, sparse data acquisition, and multidimensional decomposition. J Am Chem Soc 127:2767–2775. doi:10.1021/ja044032o

Tzeng S-R, Pai M-T, Kalodimos CG (2012) NMR studies of large protein systems. In: Shekhtman A, Burz DS (eds) Methods Mol Biol. Humana Press, Totowa, pp 133–140

Velyvis A, Schachman HK, Kay LE (2009) Assignment of Ile, Leu, and Val methyl correlations in supra-molecular systems: an application to aspartate transcarbamoylase. J Am Chem Soc 131:16534–16543. doi:10.1021/ja906978r

Velyvis A, Ruschak AM, Kay LE (2012) An economical method for production of ^{2}H, $^{13}CH_3$-threonine for solution NMR studies of large protein complexes: application to the 670 kDa proteasome. PLoS One 7:e43725. doi:10.1371/journal.pone.0043725

Xu Y, Matthews S (2013a) TROSY NMR spectroscopy of large soluble proteins. Top Curr Chem 335:97–120. doi:10.1007/128_2011_228

Xu Y, Matthews S (2013b) MAP-XSII: an improved program for the automatic assignment of methyl resonances in large proteins. J Biomol NMR 55:179–187. doi:10.1007/s10858-012-9700-z

Xu Y, Liu M, Simpson PJ et al (2009) Automated assignment in selectively methyl-labeled proteins. J Am Chem Soc 131:9480–9481. doi:10.1021/ja9020233

Yamazaki T, Lee W, Arrowsmith CH et al (1994) A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high sensitivity. J Am Chem Soc 116:11655–11666. doi:10.1021/ja00105a005