J Biomol NMR (2013) 57:237–249
DOI 10.1007/s10858-013-9784-0

ARTICLE

Differential isotope-labeling for Leu and Val residues in a protein by *E. coli* cellular expression using stereo-specifically methyl labeled amino acids

Yohei Miyanoiri · Mitsuhiro Takeda · Kosuke Okuma · Akira M. Ono · Tsutomu Terauchi · Masatsune Kainosho

Received: 23 July 2013 / Accepted: 16 September 2013 / Published online: 21 September 2013

Abstract The ^{1}H–^{13}C HMQC signals of the $^{13}CH_3$ moieties of Ile, Leu, and Val residues, in an otherwise deuterated background, exhibit narrow line-widths, and thus are useful for investigating the structures and dynamics of larger proteins. This approach, named methyl TROSY, is economical as compared to laborious methods using chemically synthesized site- and stereo-specifically isotope-labeled amino acids, such as stereo-array isotope labeling amino acids, since moderately priced, commercially available isotope-labeled α-keto acid precursors can be used to prepare the necessary protein samples. The Ile δ_1-methyls can be selectively labeled, using isotope-labeled α-ketobutyrates as precursors. However, it is still difficult to prepare a residue-selectively Leu and Val labeled protein, since these residues share a common biosynthetic intermediate, α-ketoisovalerate. Another hindering drawback in using the α-ketoisovalerate precursor is the lack of stereo-selectivity for Leu and Val methyls. Here we present a differential labeling method for Leu and Val residues, using four kinds of stereo-specifically $^{13}CH_3$-labeled [U–^{2}H;^{15}N]-leucine and -valine, which can be efficiently incorporated into a protein using *Escherichia coli* cellular expression. The method allows the differential labeling of Leu and Val residues with any combination of stereo-specifically isotope-labeled prochiral methyls. Since relatively small amounts of labeled leucine and valine are required to prepare the NMR samples; i.e., 2 and 10 mg/100 mL of culture for leucine and valine, respectively, with sufficient isotope incorporation efficiency, this approach will be a good alternative to the precursor methods. The feasibility of the method is demonstrated for 82 kDa malate synthase G.

Keywords Leu/Val differential labeling · Stereo-specifically methyl labeled Leu/Val · Methyl-methyl NOEs · Large protein

Introduction

Ile, Leu, and Val residues are the major constituents of the hydrophobic cores of proteins and protein–ligand interfaces (Janin et al. 1988). Therefore, various techniques for observing the methyl NMR signals of these residues have been developed (Metzler et al. 1996; Gardner and Kay 1997; Goto et al. 1999; Tugarinov and Kay 2004a; Tugarinov et al. 2003, 2006; Amero et al. 2009; Gans et al. 2010; Gill and Palmer 2011) and applied to explore protein structures and dynamics, especially for larger proteins and protein complexes, which are inaccessible by conventional NMR methods. The most common approach is to observe the ^{1}H–^{13}C HMQC signals for the $^{13}CH_3$ groups of Ile, Leu, and Val residues bearing protonated ^{13}C-methyls in an otherwise highly deuterated background (Tugarinov et al. 2003). In these environments, the HMQC signals have

Electronic supplementary material The online version of this article (doi:10.1007/s10858-013-9784-0) contains supplementary material, which is available to authorized users.

Y. Miyanoiri · M. Takeda · M. Kainosho (✉)
Structural Biology Research Center, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan
e-mail: kainosho@tmu.ac.jp

K. Okuma · A. M. Ono · T. Terauchi · M. Kainosho
Center for Priority Areas, Tokyo Metropolitan University, 1-1 Minami-ohsawa, Hachioji 192-0397, Japan

K. Okuma · A. M. Ono · T. Terauchi
SAIL Technologies Inc., 1-40 Suehiro-cho 1-chome, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

narrower line-widths and thus higher sensitivity, even for larger proteins. More recently, the ^{1}H–^{13}C HMQC method, which is often referred to as methyl TROSY, has been extended for observing other methyl-containing amino acids, such as alanine, threonine, and methionine (Skrynnikov et al. 2001; Gelis et al. 2007; Fischer et al. 2007; Isaacson et al. 2007; Ayala et al. 2009; Sheppard et al. 2009a, Sheppard et al. 2009b; Guo and Tugarinov 2010; Godoy-Ruiz et al. 2010; Velyvis et al. 2012). The combined use of methyl TROSY and backbone amide ^{1}H–^{15}N TROSY (Pervushin et al. 1997) has now become a standard protocol for studying larger proteins by solution NMR spectroscopy (Tugarinov et al. 2003; Tugarinov and Kay 2003, 2004a, b; Sounier et al. 2007; Amero et al. 2011; Velyvis and Kay 2013; Xu and Matthews 2013).

In the cases of branched side-chain amino acids, i.e., isoleucine, leucine, and valine, the methyl TROSY methods provide an obvious advantage for observing methyl signals, as compared to the other laborious isotope-labeling methods, such as the stereo-array isotope labeling (SAIL) method (Kainosho et al. 2006; Kainosho and Güntert 2009; Miyanoiri et al. 2012). The selectively Ile, Leu, and Val methyl labeled proteins can be efficiently prepared by standard *Escherichia coli* cellular expression using various isotope-labeled α-keto acid precursors, which are now commercially available at moderate prices (Gardner and Kay 1997; Goto et al. 1999). For example, α-ketoisovalerates with labeled methyl group(s) are used as the precursors for non-stereospecific labeling of the prochiral methyl groups of Val and Leu residues. In the case of Ile, which is produced by a different biosynthetic pathway from that for valine and leucine, methyl labeled α-ketobutyrate is used as a precursor for labeling the δ_1 methyl groups of Ile residues. A variety of regio-specifically labeled α-ketoisovalerates have been developed, including [$^{13}CH_3$, $^{13}CH_3$]- and [$^{13}CH_3$, $^{12}C^{2}H_3$]-α-ketoisovalerates, which are now used to observe the prochiral methyl signals of Leu and Val residues (Lichtenecker et al. 2004; Tugarinov and Kay 2004a, 2005; Tugarinov et al. 2006). The latter precursor is particularly useful for acquiring all of the observable inter-residue methyl–methyl NOEs for Leu and Val residues, although their intensities are 25 % of the maximum values because of this alternative methyl labeling scheme. Obviously, if the stereospecific assignment of the prochiral methyl signals is required, then a supplemental experiment, such as the biosynthetically fractional labeling method, should be used (Neri et al. 1989). Excessive crowding in the methyl regions may also be a problem, since the Val and Leu residues are both synthesized through common precursors.

In this context, further methodological developments for solving the aforementioned problems would facilitate the NMR studies of large proteins, and progress toward this goal has been reported. For example, the use of regio-specifically methyl labeled acetolactates as the precursors, in lieu of α-ketoisovalerates, allows the preparation of proteins with stereo-specifically methyl labeled Leu and Val residues (Gans et al. 2010; Mas et al. 2013). Using the acetolactates bearing a ^{13}C-labeled methyl at either the C2 or C3 position as a precursor, the Val-γ_2 (*pro-S*) and Leu-δ_2 (*pro-S*), and the Val-γ_1 (*pro-R*) and Leu-δ_1 (*pro-R*) methyls of an expressed protein are stereo-specifically labeled, respectively. A method for distinguishing Leu and Val methyl signals, using the spectral editing technique, was recently reported (Hu et al. 2012). In principle, a more straightforward and thus versatile means may be the direct incorporation of the chemically synthesized amino acids with any desirable labeling patterns into the protein to be studied, which is the basic concept of the SAIL method (Kainosho et al. 2006). The potential advantage of the stereo-specifically isotope-labeled amino acid method over the precursor method is that any of the stereo- and residue-selective combinations can be chosen for the Leu and Val methyls, such as Leu δ_2 and Val γ_1. From a practical viewpoint, however, there have been serious concerns about the overwhelming difficulties in producing SAIL proteins, because the production of a protein exclusively composed of the SAIL amino acids requires the use of cell-free expression and an expensive complete mixture of constituent SAIL amino acids (Torizawa et al. 2004; Kainosho et al. 2006; Takeda et al. 2007). However, this may not be a problem when preparing proteins with only a few selected residues labeled with SAIL amino acids. These 'selective SAIL' proteins can also be prepared by *E. coli* cellular expression, if the selected amino acid residues do not undergo metabolic scrambling during the preparation. In the case of cellular expression, the necessary amount of the SAIL amino acid can be reduced by optimizing the labeling protocols, thus substantially lowering the total cost. Indeed, we have successfully applied the cellular expression in some cases to prepare 'selective SAIL' proteins (Miyanoiri et al. 2011; Takeda et al. 2012).

Here, for the 82 kDa protein *E. coli* malate synthase G (MSG; Molina et al. 1994), we explored the differential isotope-labeling for Leu and Val residues, using site- and stereo-specifically labeled amino acids with *E. coli* cellular expression. MSG consists of 723 amino acid residues, including 70 Leu and 46 Val residues, and thus a total of 232 methyl peaks for these residues would be simultaneously observed when $^{13}CH_3$-labeled α-ketoisovalerate is used as a precursor. Therefore, the ^{1}H–^{13}C methyl TROSY spectrum was rather crowded, and this motivated us to simplify the spectra by an alternate isotope labeling method, using site- and stereo-specifically labeled Leu and Val. For that purpose, we synthesized all four isotopomers of the stereo-specifically $^{13}CH_3$-labeled [U–^{2}H; ^{15}N]-

leucine and -valine. Note that the costs to synthesize these site- and stereo-specifically labeled leucines and valines, which are designed to observe only one of the prochiral methyl signals, are substantially lower than those for the original SAIL amino acids (Kainosho et al. 2006). The synthesized Val and Leu residues were incorporated into proteins using *E. coli* cellular expression in deuterated medium. To do so, we first estimated the lowest amounts of labeled leucine and valine required for achieving sufficient incorporation efficiencies for most of the NMR studies. Obviously, one of our major concerns was the cost efficiency of the labeled amino acids to be used, as compared to that of the methods using labeled precursors. By employing the conventional and widely applicable *E. coli* cellular expression system, in lieu of the cell-free protein expression required for preparing fully SAIL-labeled proteins, and by minimizing the amounts of added labeled amino acids, we succeeded in firmly establishing an economically feasible, practical protocol for differential isotope-labeling of Leu and Val residues in a deuterated background.

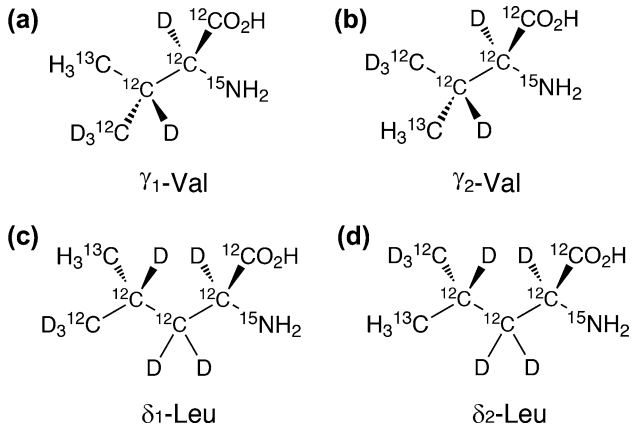


Fig. 1 Chemical structures of the stereo-specifically $^{13}CH_3$-labeled leucines and valines, bearing isotope-labeling patterns optimized for observing prochiral methyl NMR signals in large proteins. The number in parentheses for each amino acid corresponds to Schemes 1, 2, 3. **a** L-[γ_1–^{13}C;α,β,γ_2–^{2}H$_5$;α–^{15}N]-valine: "γ_1-Val" (15a); **b** L-[γ_2–^{13}C;α,β,γ_1–^{2}H$_5$;α–^{15}N] -valine: "γ_2-Val" (15b); **c** L-[δ_1–^{13}C;α,β,γ,δ_2–^{2}H$_7$;α–^{15}N]-leucine: "δ_1-Leu" (6b); **d** L-[δ_2–^{13}C;α,β,γ,δ_1–^{2}H$_7$;α–^{15}N]-leucine: "δ_2-Leu" (6a)

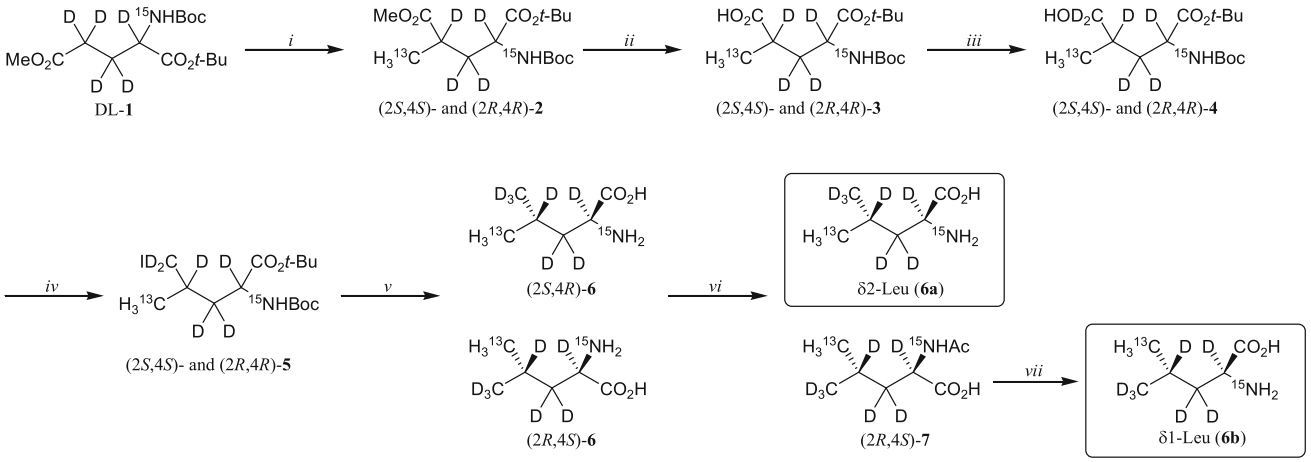
Materials and methods

Synthesis of γ_1-Val, γ_2-Val, δ_1-Leu, and δ_2-Leu

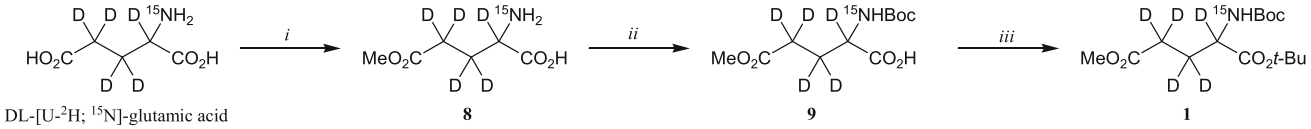
Many strategies for the stereo-selective syntheses of leucine and valine with ^{13}C-labeled methyl groups have been reported (Kelly et al. 1997; Charrier et al. 2004; Dawadi and Lugtenburg 2013). In order to synthesize all four isotopomers of the stereo-specifically $^{13}CH_3$ labeled valines; i.e., L-[γ_1–$^{13}CH_3$;α,β,γ_2–$^2H_5$;α–^{15}N]-valine ("γ_1-Val"; Fig. 1a) and L-[γ_2–$^{13}CH_3$;α,β,γ_1–$^2H_5$;α–^{15}N]-valine ("γ_2-Val"; Fig. 1b), and the stereo-specifically $^{13}CH_3$ labeled leucines; i.e., L-[δ_1–$^{13}CH_3$;α,β,γ,δ_2–$^2H_7$;α–^{15}N]-leucine ("δ_1-Leu"; Fig. 1c) and L-[δ_2–$^{13}CH_3$;α,β,γ,δ_1–$^2H_7$;α–^{15}N]-leucine ("δ_2-Leu"; Fig. 1d), we used a synthetic strategy based on a previously reported method (Oba et al. 1995). Namely, we first synthesized enantiomeric mixtures of ^{13}C-methyl labeled leucine and valine, starting from DL-[U–^{2}H;^{15}N]-glutamic acid and DL-[U–^{2}H;^{15}N]-aspartic acid, respectively, and then isolated each isotopomer by the combined use of optical resolution and racemization at the α-position.

Stereo-specifically $^{13}CH_3$-labeled leucines, namely δ_1-Leu and δ_2-Leu, were synthesized from a derivative of DL-[U–^{2}H;^{15}N]-glutamic acid (Scheme 2) via the synthetic course shown in Scheme 1. The racemic glutamate derivative **1** was treated with lithium bis(trimethylsilyl)amide (LiHMDS) and [^{13}C]-iodomethane at −78 °C to give **2** (Ardá et al. 2004). The [4-$^{13}CH_3$]-glutamate derivative **2** was then converted to labeled leucines by modifications of the methods reported by August et al. (1992) and Oba et al. (1998), as follows. The [4-$^{13}CH_3$]-glutamate **2** was treated with 1 M LiOH, to produce the [4-$^{13}CH_3$]-glutamate derivative **3**, and then was converted to the alcohol **4** by reducing the terminal carboxyl group by $NaBD_4$, via the mixed anhydride with isobutyl chloroformate. The consecutive transformation of the alcohol **4**–**5** was performed by mesylation, using methanesulfonyl chloride and triethylamine, followed by iodination with NaI. The iodide **5** was then subjected to radical deuteration by Bu_3SnD in the presence of 2,2′-azodiisobutyronitrile (AIBN). The resultant leucine derivatives were directly hydrolyzed to the L-(2*S*,4*R*)- and D-(2*R*,4*S*)-leucine **6.** Thus, after the mixture of L-(2*S*,4*R*)- **6a** and D-(2*R*,4*S*)-leucine **6b** was subjected to esterification followed by N-acetylation, enzymatic optical resolution was accomplished using *Aspergillus* acylase in the presence of $CoCl_2$ at 40 °C, to give δ_2-Leu (**6a**) in addition to the remaining D-(2*R*,4*S*)-N-acetyl leucine (**7**). In order to obtain δ_1-Leu (**6b**), the D-(2*R*,4*S*)-N-acetyl leucine (**7**) was racemized at the α-position to give a mixture of L-(2*S*,4*S*)- and D-(2*R*,4*S*)-N-acetyl leucine by a treatment with an excess of acetic anhydride, and then the mixture was subjected to enzymatic hydrolysis by the acylase to give δ_1-Leu (**6b**). The overall yields of δ_1-Leu (**6b**) and δ_2-Leu (**6a**) from DL-[U–^{2}H; ^{15}N]-glutamic acid **1** were 15 and 31 %, respectively.

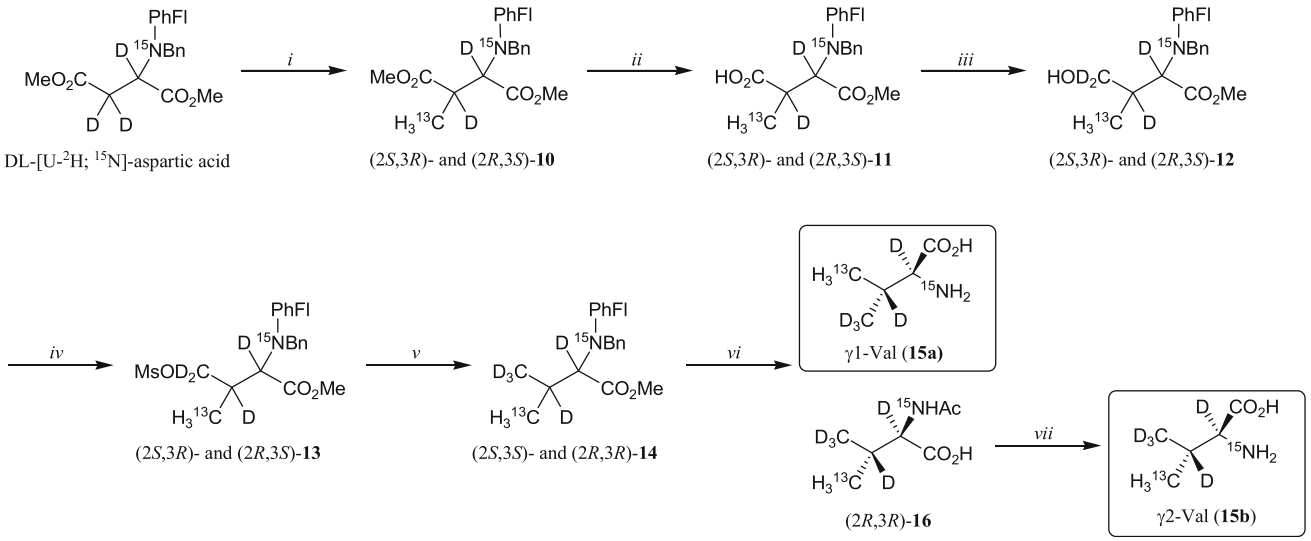
The syntheses of the two isotopomers of valine **15a** (γ_1-Val) and **15b** (γ_2-Val) were accomplished from DL-[U–^{2}H;^{15}N]-aspartic acid (Humphrey et al. 1995) by a similar strategy, using DL-[U–^{2}H;^{15}N]-aspartic acid as the starting material instead of DL-[U–^{2}H; ^{15}N]-glutamic acid (Scheme 3). The overall yields of γ_1-Val (**15a**) and γ_2-Val (**15b**), starting from DL-[U–^{2}H;^{15}N]-aspartic acid, were 14 and 8 %, respectively.



Scheme 1 Synthesis of δ_1-Leu and δ_2-Leu: *i* LiHMDS, $^{13}CH_3I$, THF, −78 °C, 69 %. *ii* LiOH, THF, H_2O. *iii* (1) iBCF, Et_3N, THF −40 °C; (2) $NaBD_4$, D_2O, 0 °C, 72 % from 2. *iv* (1) MsCl, Et_3N, DCM; (2) NaI, acetone. *v* (1) Bu_3SnD, AIBN, benzene, reflux; (2) 1M HCl, 110 °C; (3) Dowex 50W-X8. *vi* (1) EtOH, HCl; (2) AcCl, Et_3N, DCM; (3) LiOH; (4) Acylase, $CoCl_2$, 40 °C; (5) Dowex 50W-X8, 95 % from 4. *vii* (1) NaOD, Ac_2O; (2) Acylase, $CoCl_2$, 40 °C; (3) Dowex 50W-X8



Scheme 2 Synthesis of the glutamic acid derivative as an intermediate for the synthesis of δ_1-Leu and δ_2-Leu: *i* HCl/MeOH, 91 %. *ii* Boc_2O, Na_2CO_3. *iii* DMF-dineopentylacetal, *t*-BuOH, 72 % from 8



Scheme 3 Synthesis of γ_1-Val and γ_2-Val: *i* LiHMDS, $^{13}CH_3I$, THF, −78 °C. *ii* LiOH, THF, H_2O. *iii* (1) iBCF, Et_3N, THF, −40 °C; (2) $NaBD_4$, D_2O, 0 °C. *iv* (1) MsCl, Et_3N, DCM. *v* Zn, NaI, diglyme, D_2O. *vi* (1) Pd/C, H_2, TFA; (2) AcCl, Et_3N, DCM; (3) LiOH; (4) Acylase, $CoCl_2$, 40 °C; (5) Dowex 50W-X8. *vii* (1) NaOD, Ac_2O; (2) Acylase, $CoCl_2$, 40 °C; (3) Dowex 50W-X8

Preparation of [U–^{2}H;^{15}N]-MSG with stereo-specifically $^{13}CH_3$ labeled leucines and valines using *E. coli* cellular expression

The cellular expression of MSG proteins was performed according to the protocol described previously (Tugarinov et al. 2002), with some minor modifications. The *E. coli* BL21(DE3) pLys S cells, which were transformed with pMSG-B encoding the MSG protein with a His-tag at its C-terminus (Howard et al. 2000), were cultured at 37 °C in 100 ml of M9 medium in D_2O (CIL), containing 7.0 g/L Na_2HPO_4, 3.0 g/L KH_2PO_4, 0.50 g/L NaCl, 1.0 g/L

[U–^{2}H]-glucose (CIL), 1.0 g/L $^{15}NH_4Cl$, 20 mg/L thiamine hydrochloride, 20 mg/L (+)-biotin, 1.6 mg/L $FeCl_3$, 0.24 g/L $MgSO_4$, 6.3 mg/L $MnCl_2$, 11 mg/L $CaCl_2$, 0.10 mg/L folic acid, 0.10 mg/L choline chloride, 0.1 mg/L nicotinamide, 0.1 mg/L D-pantothenic acid, 10 μg/L riboflavin, 0.1 mg/L pyridoxal hydrochloride, and 0.2 mg/L *myo*-inositol.

In order to reduce the intracellular levels of unlabeled leucine and valine at the time of isopropyl-β-D-thiogalactopyranoside (IPTG) induction for protein production, we found that culturing the cells with one-third of the total amount of added labeled leucine and valine was effective to increase the incorporation efficiencies of leucine and valine, to some extent. For example, if we wanted to label MSG with a total of 10 mg of γ_1-Val or γ_2-Val, then we added 3.3 mg of the labeled valine to 100 mL of deuterated M9 medium at the start of cell cultivation. The remaining 6.7 mg of labeled valine was added when the OD_{600} reached about 0.4, and then the IPTG induction was started. We also prepared MSG selectively labeled with valine and leucine using biosynthetic precursors, for the sake of comparison between our method using site- and stereo-specifically labeled amino acids. For that purpose, 12 mg of [3-$^{13}CH_3$; 3,4,4,4-$^{2}H_4$]-α-ketoisovaleric acid sodium salt (CIL) or 30 mg of [2-$^{13}CH_3$; 4,4,4-$^{2}H_3$]-acetolactate (Gans et al. 2010) was added to 100 ml of deuterated M9 medium when the OD_{600} reached about 0.3. When the OD_{600} reached 0.4 about 1 h later, IPTG was added to a final concentration of 1 mM, and the growth was continued for 8 h at 37 °C with gentle shaking (150 rpm). The *E. coli* cells were then harvested by centrifugation and processed for MSG preparation (Tugarinov et al. 2002).

Quantitative analyses of the incorporation efficiency of Val and Leu residues in MSG using mass spectrometry

The incorporation efficiencies of the Leu and Val residues in MSG, prepared with various amounts of added isotopically labeled amino acids, were quantitatively estimated by a GC/MS analysis on a GC/MS QP 2010 spectrometer (Shimadzu). MSG samples prepared from cells grown in the deuterated M9 medium supplemented with "unlabeled" leucine and/or valine, under the same conditions as described above, were conveniently used for this purpose. The expressed MSG samples were fully deuterated, except for the Leu and/or Val residues, which were partially deuterated according to the incorporation efficiencies. A 100 μL aliquot of about 200 μM MSG solution was mixed with 400 μL of 4 M methanesulfonic acid, and was heated in a sealed tube at 160 °C for 45 min. The hydrolyzed amino acid mixture was derivatized using an EZ:Faast amino acid analysis kit (Phenomenex), according to the manufacturer's recommended protocol. The resultant amino acid derivatives were subjected to the GC/MS analysis. The detailed procedures for the mass spectrometric analysis will be published elsewhere (Terauchi et al. to be published).

NMR spectroscopy

MSG was purified from the *E. coli* cell cultures, by anion-exchange chromatography after the Ni–NTA column purification, and then the protein was unfolded and refolded to ensure that the backbone amide proton back exchange was complete (Tugarinov et al. 2002). The refolded labeled MSG was dialyzed against NMR buffer, containing 20 mM sodium phosphate, 20 mM magnesium chloride, 5 mM deuterated DL-1,4-dithiothreitol (d_{10}-DTT) and 1 mM sodium azide, at pH 7.1. The NMR buffer was prepared with 99 % H_2O/1 % D_2O. We usually obtained about 2.5 mg of purified MSG from a 100 ml deuterated M9 culture, which was sufficient to make one NMR sample (0.15 mM/200 μL) in a slotted Shigemi tube (Shigemi Co.). These salt-tolerant sample tubes require less sample volume and provide higher signal-to-noise ratios for aqueous sample solutions, as compared to conventional Shigemi tubes, when the quasi-rectangular slot is properly positioned against the rf coil of the cryoprobe, using the sample positioning unit (Takeda et al. 2011). The NMR measurements were performed at 37 °C using an Avance900 spectrometer, equipped with a TCI cryoprobe and the sample positioning unit (Bruker Biospin).

In the 2D ^{1}H–^{13}C methyl TROSY experiments (Tugarinov et al. 2003), the data size and spectral width were 256 (t_1) × 2,048 (t_2) and 3,200 Hz (ω_1, ^{13}C) × 14,400 Hz (ω_2 ^{1}H), respectively. The ^{1}H and ^{13}C carrier frequencies were 4.7 and 20 ppm, respectively. The number of scans/FID was 32. The repetition time was 2 s. In the ^{13}C-edited 3D NOESY-HMQC experiment, the data size and spectral width were 256 (t_1) × 24 (t_2) × 2048 (t_3) and 14,400 Hz (ω_1, ^{1}H) × 3,200 Hz (ω_2, ^{13}C) × 14,400 Hz (ω_3, ^{1}H), respectively. The ^{1}H and ^{13}C carrier frequencies were 4.7 and 20 ppm, respectively. The number of scans/FID was 16. The NOE mixing time and repetition time were 300 ms and 2 s, respectively. All NMR spectra were processed with the TopSpin software, version 3.1 (Bruker Biospin).

Results and discussion

Optimization of the individual incorporation efficiencies for Leu and Val residues in a protein using *E. coli* cellular expression

Two major concerns in differential isotope-labeling for Leu and Val residues in a protein using the cellular expression

system are the labeling efficiency and the residue-selectivity. To estimate the minimum amounts of leucine and valine needed to achieve acceptable incorporation efficiencies (>80 %) and to monitor the isotope scrambling during protein expression, we used the GC/MS method to quantitatively analyze the isotopomer distributions of leucine and valine isolated from the acid hydrolyzates of MSG samples. These proteins were produced in deuterated M9 media containing different amounts of unlabeled, and thus protonated, leucine and valine. Under these culturing conditions, the expressed MSG contains both highly deuterated (endogenous) and highly protonated (exogenous) Leu and Val residues. These two types of Leu and Val residues from different origins can be easily differentiated by the relative intensities of the appropriate fragment ions in the mass spectra. Therefore, the incorporation efficiencies could be quantitatively determined for the Leu and Val residues in the MSG sample, as a function of the amounts of leucine or valine added to the cultures, as illustrated in Fig. 2a and b (See also Table S1). In the case of Leu-selective labeling, when only unlabeled leucine was added to the culture medium, we found that 20 mg/L of leucine resulted in an over 90 % incorporation efficiency, and its metabolic scrambling to other amino acids was not detected (Fig. 2a, Table S1).

On the other hand, in the case of Val-selective labeling, the incorporation efficiency of exogenous valine was comparatively low, and its conversion to leucine was clearly observed. For example, the incorporation efficiency was ~70 %, when the culture was supplemented with 25 mg/L of valine. Even when the supplemented amount was increased to 100 mg/L, the incorporation efficiency was still around 80 % (Fig. 2b, Table S1). This marked difference in incorporation efficiencies between Leu and Val residues can be explained by their biosynthetic paths in the *E. coli* cells. According to the well-established amino acid biosynthetic pathways in microorganisms, valine is synthesized from two pyruvate molecules through four enzymatic steps (Fig. 2c). Since the intracellular pyruvate concentration should be relatively high in the cells, it is not possible to completely inhibit the biosynthetic pathway simply by increasing the concentration of exogenous valine up to 100 mg/L. This explanation is reasonable, since the ^{1}H–^{13}C cross peaks were observed for Val γ_1 and γ_2 methyls in the ^{2}H-decoupled, {^{2}H}, ^{1}H–^{13}C HMQC spectrum of MSG produced in the deuterated M9 medium containing [U–^{13}C]-glucose, 100 mg/L [U–^{2}H]-Val, and 20 mg/L [U–^{2}H]-Leu (Fig. S1b). In the biosynthetic pathways of *E. coli*, the methyl groups of amino acids are derived from the methyl group of pyruvate, which originated from either a CH or a CH_2 moiety of glucose. In this condition, the major isotopomer in the methyl groups of endogenous amino acids is the $^{13}C^{1}HD_2$ isotopomer (Rosen et al. 1996; Shekhtman et al. 2002; Otten et al. 2010; Guo and Tugarinov 2010). As described in detail in the caption of Figure S1, the Val residues were shown to have "$^{13}C^{1}HD_2$" methyls; therefore, they were from the partially deuterated [U–^{13}C]-valine biosynthetically derived from [U–^{13}C]-glucose in the deuterated medium. On the other hand, the Leu methyl signals were not observed, since the conversion from intracellular valine to leucine was strongly suppressed by the feed-back inhibition by [U–^{2}H]-Leu in this MSG preparation (Fig. 2c; Table S1; Fig. S1b). Due to the effective feed-back inhibition of the valine biosynthesis by leucine, the incorporation efficiency of Val residues using exogenous labeled valine was slightly increased by adding leucine to the culture medium. For example, the incorporation efficiency of Val residues using 100 mg/L of exogenous valine was increased up to ~83 % in the presence of 20 mg/L of leucine (Table S1).

Attempts to further increase the incorporation efficiency for valine by suppressing the de novo valine synthesis in *E. coli* cells, using acetolactate synthase inhibitors, were not very successful (Table S1). Therefore, in order to selectively label Val residues, it was necessary to add a sufficient amount (~20 mg/L) of unlabeled (or deuterated) leucine to the culture medium; otherwise, the Leu residues are simultaneously labeled (Table S1). However, in the case of combinatorial Val/Leu labeling using various isotopically labeled valines and leucines, the undesired metabolic conversion from valine to leucine can be concomitantly suppressed by the co-existing leucine.

Differential labeling of Leu and Val residues in MSG using stereo-specifically $^{13}CH_3$ labeled amino acids

In order to observe the prochiral methyl signals of Leu and Val residues in large proteins most efficiently, we chemically synthesized the following stereo-specifically ^{13}C-methyl labeled valines and leucines: L-[γ_1–^{13}C;α,β,γ_2–$^{2}H_5$;α–^{15}N]-valine ("γ_1-Val"; Fig. 1a), L-[γ_2–^{13}C;α,β,γ_1–$^{2}H_5$;α–^{15}N]-valine ("γ_2-Val"; Fig. 1b), L-[δ_1–^{13}C;$\alpha,\beta,\gamma,\delta_2$–$^{2}H_7$;$\alpha$–$^{15}N$]-leucine ("$\delta_1$-Leu"; Fig. 1c), and L-[$\delta_2$–$^{13}C$;$\alpha,\beta,\gamma,\delta_1$–$^{2}H_7$;$\alpha$–$^{15}N$]-leucine ("$\delta_2$-Leu"; Fig. 1d). These amino acids have an isolated $^{13}CH_3$ group with a designated stereochemistry, and therefore only one of the prochiral methyl signals can be observed with very high-sensitivity, when they are efficiently incorporated into a protein. Since all of the protons, except for the ^{13}C-labeled site, are fully deuterated in these amino acids, NMR measurements for the labeled Leu and Val residues in a protein do not suffer from scalar and dipolar interactions between intra-residue spins. The benefits of these optimized labeling patterns are predicated upon their incorporation in a deuterated background. The ^{15}N labeling is not necessary to observe the Leu and Val

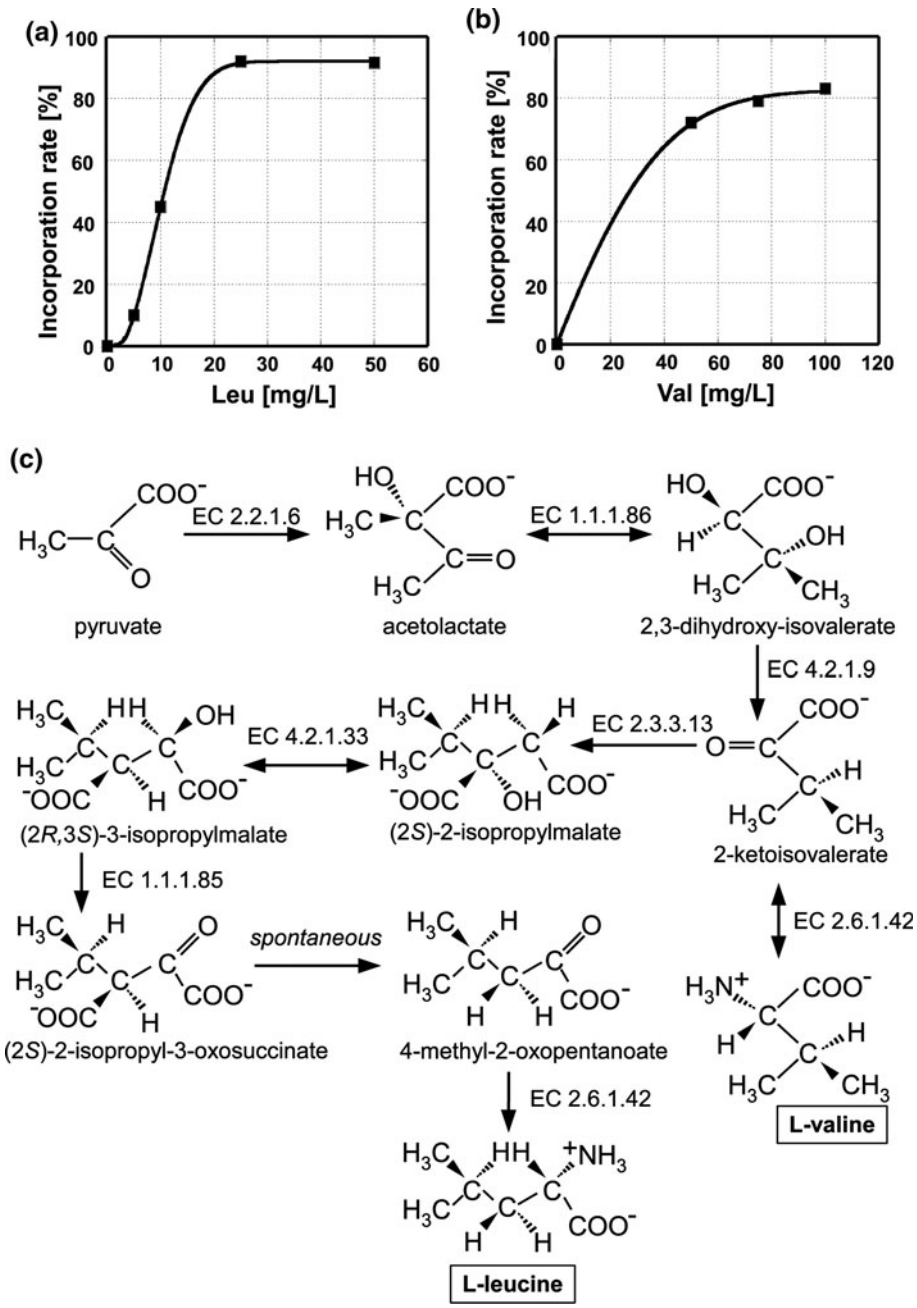


Fig. 2 The incorporation rates of "exogenous" leucine and valine determined for MSG samples isolated from *E. coli* cells cultured in deuterated M9 medium, supplemented with various amounts of *unlabeled* leucine and valine. **a** The incorporation rate changes for the Leu residues by adding *unlabeled* leucine in the range of 5–50 mg/L. **b** The incorporation rate changes for the Val residues by adding *unlabeled* valine in the range of 50–100 mg/L, together with 20 mg/L of [U–^{2}H]-leucine. The incorporation rates were quantitatively determined by mass-spectrometry, by measuring the relative intensities between highly deuterated "endogenous" leucine or valine, and highly protonated "exogenous" *unlabeled* leucine or valine, by using a few relevant fragment ions to characterize the origins of leucine and valine in the isolated MSG. Details of the mass-spectrometric analyses will be published elsewhere (Terauchi et al. to be published). **c** Biosynthetic pathways of leucine and valine in *E. coli* cells. The EC numbers indicate the enzymes responsible for each enzymatic reaction pathway. The biosynthetic intermediates and the metabolic pathways for leucine and valine are available at the following web-site of the International Union of Biochemistry and Molecular Biology (http://www.chem.qmul.ac.uk/iubmb/)

methyls, but it may be useful for various future applications of these labeled amino acids.

We prepared MSG samples selectively labeled with either γ_1-Val, γ_2-Val, δ_1-Leu, or δ_2-Leu in a deuterated background. In order to maximize the incorporation efficiencies while using the smallest possible amounts of labeled amino acids, we followed the procedures described above. Namely, 10 mg of either γ_1-Val or γ_2-Val was used in 100 mL deuterated M9 medium, which was supplemented with 2 mg of [U–^{2}H]-leucine to suppress the intracellular conversion from valine to leucine and thus increase the incorporation efficiencies for γ_1-Val and γ_2-Val. Each of these cultures yielded 200 µL of a 0.15 mM NMR sample, an ideal volume for use in a slotted Shigemi tube (Takeda et al. 2011), and the 46 Val residues in MSG gave well-dispersed ^{1}H–^{13}C HMQC signals for Val γ_1 and Val γ_2, as shown in Fig. 3a, b, respectively. In the case of Leu-selective labeling, 2 mg of either δ_1-Leu or δ_2-Leu was added to 100 mL of deuterated medium, to yield residue- and stereo-specifically Leu labeled MSG samples. These samples gave well-dispersed ^{1}H–^{13}C HMQC signals for 70 Leu residues in MSG, as shown in Fig. 3c for Leu

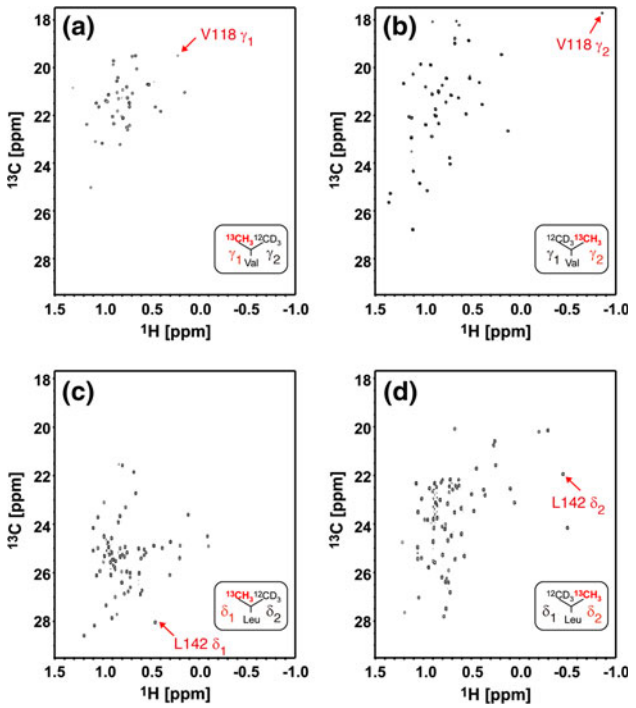


Fig. 3 900 MHz methyl-TROSY spectra of the residue-selectively Leu- or Val-labeled MSG samples, using stereo-specifically $^{13}CH_3$-labeled leucines or valines. Labeled MSG samples, prepared by the protocol described in the section "Materials and methods", were dissolved in a 200 μL volume. All spectra were obtained for 200 μL of a 0.15 mM solution at pH 7.1, 37 °C. **a** "γ_1-Val" labeled [U–^{2}H,^{15}N]-MSG; **b** "γ_2-Val" labeled [U–^{2}H,^{15}N]-MSG; **c** "δ_1-Leu" labeled [U–^{2}H,^{15}N]-MSG; **d** "δ_2-Leu" labeled [U–^{2}H,^{15}N]-MSG. The labeling patterns of the Leu and Val methyl groups are shown in the *box* for each spectrum. Signal assignments for V118 γ_1 (*pro-R*), γ_2 (*pro-S*) and L142 δ_1 (*pro-R*), δ_2 (*pro-S*) methyls are shown in each spectrum (Tugarinov and Kay 2003)

δ_1, and Fig. 3d for Leu δ_2. As expected, there was no interconversion between Val γ_1 and Val γ_2, or Leu δ_1 and Leu δ_2.

These four stereo- and residue-selectively methyl labeled proteins allowed us to observe the methyl–methyl NOEs with fourfold higher sensitivities for either the Val γ_1–Val γ_1, Val γ_2–Val γ_2, Leu δ_1–Leu δ_1, or Leu δ_2–Leu δ_2 pairs in a protein, as compared to those for the protein labeled with [3-$^{13}CH_3$;3,4,4,4-$^2H_4$]-α-ketoisovalerate as a precursor. However, the inter-residue NOEs for the prochiral methyls belonging to Leu, or the Val residues with different stereochemistries, namely, Val γ_1–Val γ_2, or Leu δ_1–Leu δ_2 pairs, cannot be measured with these residue- and stereo-selectively methyl labeled proteins. For that purpose, an equimolar mixture of either γ_1-Val and γ_2-Val (a total of 100 mg/L labeled valine mixture and 20 mg/L unlabeled or deuterated leucine), or δ_1-Leu and δ_2-Leu (a total of 2 mg/L of labeled leucine mixture) should be supplemented to the medium, although the incorporation efficiencies for each of the prochiral methyls are decreased to half of those of the stereo-specifically labeled proteins. This type of information can also be obtained with the same sensitivity by using [3-$^{13}CH_3$;3,4,4,4-$^2H_4$]-α-ketoisovalerate as the precursor. Although one labeled protein sample prepared by the precursor method can afford all possible combinations for the inter-residue methyl–methyl NOEs associated with Leu and Val residues, in practice, the excessive signal overlapping may hamper the acquisition of all of the NOE data (Table S2). Note that Leu-selective isotope-labeling cannot be achieved by using α-ketoisovalerate or acetolactate precursors.

Combinatorial labeling for the Leu and Val residues in MSG using stereo-specifically $^{13}CH_3$ labeled Leu and Val

We mentioned above that the prochiral methyls of Leu and Val residues in a protein are alternatively labeled if [3-$^{13}CH_3$; 3,4,4,4-$^2H_4$]-α-ketoisovalerate is used as a precursor. Since one of the prochiral methyls in the Leu and Val residues prepared using this precursor is always $^{12}C^2H_3$, there will be no intra-residue methyl–methyl NOEs, and thus only inter-residue NOEs among the Leu and Val methyls can be observed. This feature makes [3-$^{13}CH_3$; 3,4,4,4-$^2H_4$]-α-ketoisovalerate a very useful precursor for methyl labeling of Leu and Val residues. However, because the incorporation efficiency for each methyl is less than 50 %, the intensities for the inter-residue methyl–methyl NOEs between Leu and Val residues become ~25 % of the maximal value (achievable in [$^{13}CH_3$,$^{13}CH_3$]-Leu/Val labeled samples). The situation can be improved by using selectively ^{13}C-labeled acetolactates (Gans et al. 2010; Mas et al. 2013), which are the precursors for the valine-leucine biosynthetic pathways (Fig. 2c). Since the Val γ_1 and γ_2 methyls are biosynthetically transferred into the Leu δ1 and δ2 methyls, respectively, the regio-selectively ^{13}C-labeled acetolactate precursors are also useful to prepare proteins with simultaneously and stereo-specifically methyl-labeled Val and Leu residues. Therefore, by choosing [2-$^{13}CH_3$;4,4,4-$^2H_3$]- or [2-$^{12}C^2H_3$;1,2,3,4-$^{13}C_4$]-acetolactate, the Val γ_2 and Leu δ_2, and the Val γ_1 and Leu δ_1 methyls, can be stereo-specifically labeled, respectively (Gans et al. 2010; Mas et al. 2013). These Leu/Val doubly-labeled proteins are useful to measure the inter-residue methyl–methyl NOEs, for either the Val γ_1–Leu δ_1 or Val γ_2–Leu δ_2 pairs in proteins, with higher sensitivities as compared to those obtained for a protein prepared using α-ketoisovalerate precursors. However, it should be noted that the Val and Leu methyls labeled with acetolactate precursors have the same stereochemistry, and therefore one cannot use them for observing the methyl–methyl NOEs between the Leu and Val residues with different stereochemistries.

On the other hand, any combinatorial labeling patterns are possible by the cellular expression method using stereo-specifically labeled leucine and valine, if the metabolic scrambling reactions for the labeled amino acids are suppressed. We successfully prepared MSG samples bearing all four combinations of stereo-specifically methyl-labeled Leu and Val residues and measured their methyl TROSY spectra, as shown in Fig. 4: Val γ_1–Leu δ_1 (Fig. 4a), Val γ_1–Leu δ_2 (Fig. 4c), Val γ_2–Leu δ_1 (Fig. 4d), and Val γ_2–Leu δ_2 (Fig. 4b). Among these four Leu–Val combinatorially labeled proteins, the Val γ_1–Leu δ_2 and Val γ_2–Leu δ_1 pairs can only be prepared by using stereo-specifically labeled leucines. As mentioned above, each NMR sample solution was prepared from a 100 mL culture supplemented with 10 mg of labeled valine and 2 mg of labeled leucine in a deuterated background. As illustrated for the L142 and V118 methyl signals, there is no stereochemical conversion for the prochiral methyls, as expected. These Leu–Val combinatorially labeled proteins provided the opportunities for higher-sensitivity measurements of the inter-methyl NOEs for Leu and Val residues with different stereochemistries, as described in the next chapter.

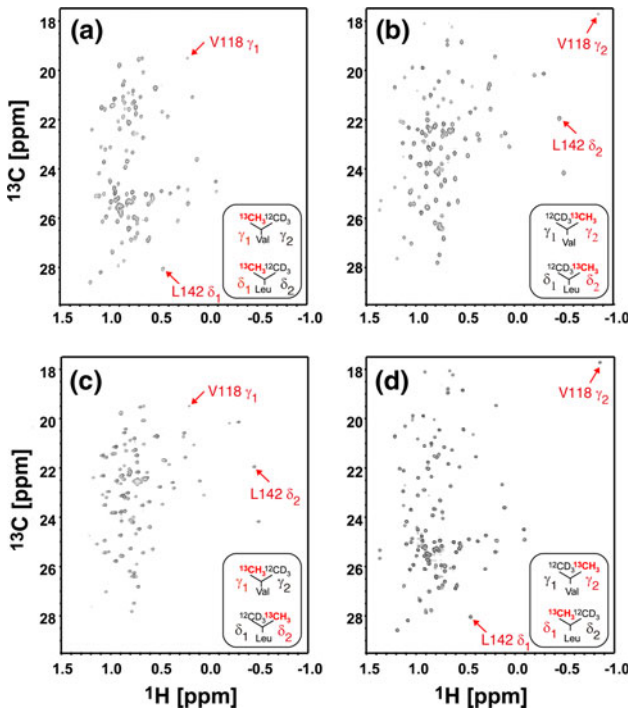


Fig. 4 900 MHz methyl-TROSY spectra of the combinatorially Leu/Val-labeled MSG samples, using stereo-specifically $^{13}CH_3$-labeled leucines and valines. Combinatorially Leu/Val labeled MSG samples, prepared by the protocol described in the section "Materials and methods", were dissolved in a 200 µL volume. All spectra were obtained for 200 µL of a 0.15 mM solution at pH 7.1, 37 °C. **a** "γ_1-Val + δ_1-Leu" labeled [U–^{2}H,^{15}N]-MSG; **b** "γ_2-Val + δ_2-Leu" labeled [U–^{2}H,^{15}N]-MSG; **c** "γ_1-Val + δ_2-Leu" labeled [U–^{2}H,^{15}N]-MSG; **d** "γ_2-Val + δ_1-Leu" labeled [U–^{2}H,^{15}N]-MSG. The combinations of the labeling patterns of the Leu/Val methyl groups are shown in the *box* for each spectrum. Signal assignments for V118 γ_1 (*pro-R*), γ_2 (*pro-S*) and L142 δ_1 (*pro-R*), δ_2 (*pro-S*) methyls are shown in each spectrum (Tugarinov and Kay 2003)

The incorporation efficiencies for the Val and Leu residues were analyzed by mass spectrometry and were found to be around 83 and 92 %, respectively; therefore, the signal intensities of the Leu residues are higher than those of the Val residues (Fig. 4; Table S1). We observed a slightly lower incorporation efficiency for the Val residues in MSG, even with a fivefold higher concentration of exogenous labeled valine in the culture medium (100 mg/L), as compared to that of labeled leucine (20 mg/L). Therefore, we suspected that the exogenous labeled valine becomes diluted in the cells by the endogenous unlabeled valine, even though the synthetic pathways from the acetolactate intermediate to valine are inhibited by the addition of labeled valine and leucine to the culture medium. We confirmed that this was actually the case by obtaining the {^{2}H} ^{1}H–^{13}C HMQC spectrum of MSG prepared from the *E. coli* cells cultured in deuterated M9 medium composed of [U–^{13}C]-glucose, in lieu of unlabeled glucose, and [U–^{2}H]-valine and -leucine. As shown in Figure S1 (b), the γ_1 and γ_2 "$^{13}C^1HD_2$" methyl signals from to the ^{13}C-labeled Val residues were clearly observed in the high-field ^{1}H-chemical range. These residues showed approximately 0.6 ppm higher ^{13}C-chemical shifts of their $^{13}C^1H_3$ signals, due to the secondary isotope shift caused by the directly bonded deuteron. In addition, the much stronger "$^{13}C^1HD_2$" methyl signals for Ile γ_2 and δ_1 and Ala β were observed, but no "$^{13}C^1HD_2$" methyl signals were detected for the Leu methyls. This observation can be perfectly explained by the mass spectrometric data in Table S1, with details provided in the caption of Figure S1.

Comparisons of the methyl–methyl NOEs for the Leu- and Val-residues in MSGs prepared using labeled α-keto acid precursors and stereo-specifically $^{13}CH_3$-labeled amino acids

As described above, *E. coli* cellular expression using stereo-specifically $^{13}CH_3$-labeled leucines and valines, in lieu of isotope-labeled keto acid precursors, allowed us to prepare proteins bearing labeled Leu/Val methyl groups with any designated stereochemical combinations. It is especially important to have all four combinatorially Leu/Val $^{13}CH_3$-labeled proteins available for analyses of large proteins, since the methyl–methyl NOEs associated with the Leu and Val methyls are extremely valuable sources of structural information. In principle, all of the methyl–methyl NOEs among the prochiral methyl protons of the Leu and Val residues can be measured with a single sample prepared using [3-$^{13}CH_3$;3,4,4,4-$^2H_4$]-α-ketoisovalerate in deuterated medium, as illustrated in Fig. 5a for a selected

2D [^{1}H,^{1}H]-plane of the 3D ^{13}C-edited NOESY-HMQC spectrum, obtained for the MSG prepared using this precursor. Figure 5b shows the same spectral region observed for the MSG simultaneously labeled with 2 mg/100 mL of L-[δ_2–^{13}C;$\alpha,\beta,\gamma,\delta_1$–^{2}H$_7$;$\alpha$–^{15}N]-leucine, "$\delta_2$-Leu", and 10 mg/100 mL of L-[γ_1–^{13}C;α,β,γ_2–^{2}H$_5$;α–^{15}N]-valine, "γ_1-Val", in the deuterated medium. This combinatorially Leu- and Val-labeled MSG in a deuterated background has approximately twofold higher concentrations of the Leu δ_2 and Val γ_1 ^{13}CH$_3$ groups, while the Leu δ_1 and Val γ_2 methyls are fully deuterated. Consequently, the NOE cross peaks in Fig. 5b showed much stronger signals. There are a few extra NOEs in Fig. 5a, as compared to those observed in Fig. 5b. These extra NOEs are from the spatially close Leu δ_1 and Val γ_2 methyls, leading to more crowding in Fig. 5a. In order to compare the NOE signal relative intensities versus the inter-methyl distances, the partial assignments for the NOEs commonly observed in the two spectra are shown in Fig. 5a, b. The relative intensities of the NOE cross-peaks at a mixing time of 300 ms were measured for Fig. 5b, and were compared to the H–H distances between the relevant methyl groups, according to the crystal structure (PDB #1d8c, Howard et al. 2000): L230 δ_2-V217 γ_1 (*relative NOE intensity* 9.5_6; *H–H distance* 3.7_4 Å), L230 δ_2–L329 δ_2 (2.8_3; 5.4_2 Å), L25 δ_2–L413 δ_2 (1.9_1; 5.9_5 Å), and V365 γ_1–L413 δ_2 (2.4_4; 5.0_3 Å). The NOE intensities are normalized according to the incorporation efficiencies for Leu δ (92 %) and Val γ (83 %), as shown in Table S1. The protein concentration in both samples was 0.3 mM, and the other experimental conditions were the same, as described in the *Materials and Methods*. Therefore, using combinatorially Leu- and Val-methyl labeled samples, one can obtain NOEs between methyls with H–H distances of about 6 Å. However, such long-range inter-methyl NOEs are much weaker for a protein labeled with the [3-^{13}CH$_3$;3,4,4,4-^{2}H$_4$]-α-ketoisovalerate precursor, as shown in Fig. 5a. As pointed out previously, by using [2-^{13}CH$_3$;4,4,4-^{2}H$_3$]- or [2-^{12}C^{2}H$_3$;1,2,3,4-^{13}C$_4$]-acetolactate, stereo-specific combinatorial labeling of the Val γ_2/Leu δ_2, or Val γ_1/Leu δ_1 methyls, respectively, can be achieved (Gans et al. 2010; Mas et al. 2013). These proteins are useful to obtain inter-methyl NOEs for the Val γ_1–Leu δ_1 or Val γ_2–Leu δ_2 pairs in proteins, with higher sensitivities. While these acetolactate precursors cannot be used for labeling Leu and Val methyls with different stereochemistries, such proteins can be prepared by the addition of stereo-specifically ^{13}CH$_3$-labeled leucine to the labeled acetolactate with different stereochemistry.

In some cases, the Leu/Val methyl ^{1}H–^{13}C signals incidentally overlapped with each other, even at a very high-field. An example is shown in Fig. 5c, for the MSG prepared using the [3-^{13}CH$_3$;3,4,4,4-^{2}H$_4$]-α-ketoisovalerate precursor. In this sample, V118 γ_1 and V193 γ_2 appeared as an unresolved single cross-peak in the 2D methyl TROSY spectrum (Fig. 5c). This peak was shown to be composed of two closely spaced peaks, by measuring the methyl TROSY spectra of the γ_1 Val-labeled and γ_2 Val-labeled MSG. The overlaid signals for the V118 γ_1 (red) and V193

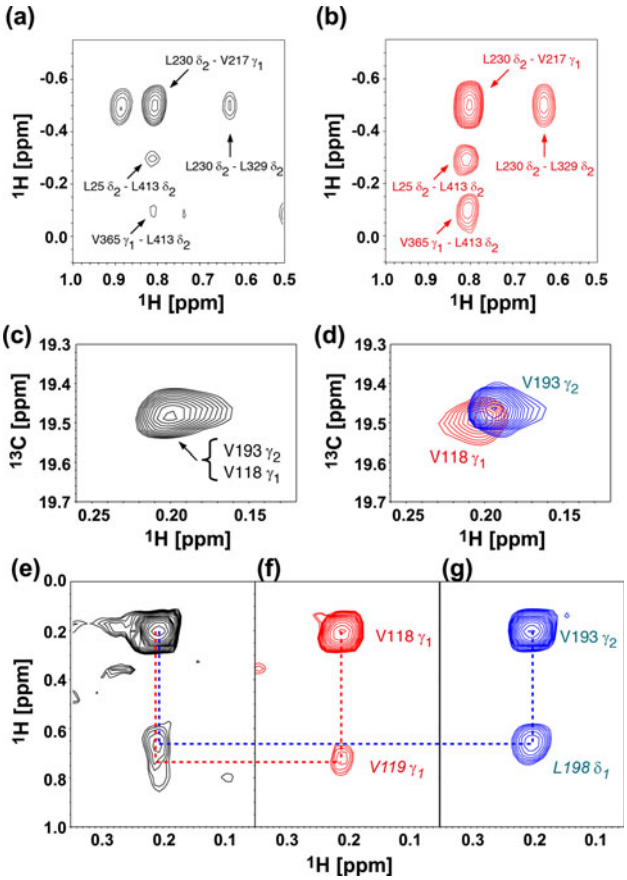


Fig. 5 Comparison of the Leu/Val methyl–methyl NOEs in MSG samples prepared by [3-^{13}CH$_3$; 3,4,4,4-^{2}H$_4$]-α-ketoisovalerate and by stereo-specifically ^{13}CH$_3$-labeled leucines and valines. **a** A 2D [^{1}H,^{1}H] plane from the 3D ^{13}C edited NOESY-HMQC spectrum, at the ^{13}C position of 22.12 ppm, observed for a 0.3 mM solution of MSG prepared by using the [3-^{13}CH$_3$; 3,4,4,4-^{2}H$_4$]-α-ketoisovalerate precursor in deuterated M9 medium; **b** A 2D [^{1}H,^{1}H] plane from the 3D ^{13}C edited NOESY-HMQC spectrum, at the ^{13}C position of 22.12 ppm, observed for a 0.3 mM solution of [U–^{2}H,^{15}N]-MSG combinatorially labeled with an equimolar mixture of "γ_1-Val" and "δ_2-Leu"; **c** The overlapped methyl TROSY signals of V193 γ_2 and V118 γ_1, observed for MSG labeled with the [3-^{13}CH$_3$; 3,4,4,4-^{2}H$_4$]-α-ketoisovalerate precursor; **d** Overlaid methyl TROSY signals for V193 γ_2 (*blue*) and V118 γ_1 (*red*), observed for the MSG samples stereo-specifically labeled with either γ_2-Val or γ_1-Val; **e** Unresolved NOE signals obtained from the overlapped V193 γ_2 and V118 γ_1 methyl TROSY signals for MSG labeled with the [3-^{13}CH$_3$; 3,4,4,4-^{2}H$_4$]-α-ketoisovalerate; **f**, **g** The NOEs starting from V118 γ_1 to *V119* γ_1 (*red*), and from V193 γ_2 to *L198* δ_1 (*blue*) were clearly observed for the MSG samples combinatorially labeled with "γ_1-Val + δ_2-Leu", and with "γ_2-Val + δ_1-Leu", respectively. The results revealed that the unresolved NOE in **e** is actually two overlapped signals

γ_2 (blue) revealed that the chemical shift differences of these peaks are very small for both the ^{1}H ($\sim$0.01 ppm) and ^{13}C dimensions ($\sim$0.03 ppm) (Fig. 5d). These incidentally overlapped signals might lead to erratic NOE assignments. Since the chemical shifts of V118 γ_1 and V193 γ_2 are overlapped, the 2D [^{1}H,^{1}H] plane of the 3D ^{13}C-edited NOESY-HMQC spectrum containing these signals showed unresolved NOE cross peaks, which could not be assigned. However, the same planes observed for the Val γ_1–Leu δ_2 and Val γ_2–Leu δ_1 combinatorially labeled MSG samples showed a weak NOE between V118 γ_1 and V119 γ_1 (*relative NOE intensity* 1.9_4; *H–H distance* 6.2_9 Å), Fig. 5f, and a strong NOE between V193 γ_2 and L198 δ_1 (3.2_8; 4.6_1 Å), Fig. 5g, respectively. In a similar manner to the NOE data in Fig. 5b, the relative NOE intensities are normalized by the incorporation efficiency for Leu δ (92 %) and Val γ (83 %). The ^{1}H-chemical shift difference between V119 γ_1 and L198 δ_1 was found to be less than 0.1 ppm, and thus it cannot be distinguished without using these two combinatorial Leu–Val $^{13}CH_3$-labeled MSGs with different stereochemistries.

Conclusions

We have developed a practical protocol for the differential labeling of Leu and Val residues in a protein, using the conventional *E. coli* cellular expression system with chemically synthesized L-[δ_1–^{13}C;α,β,γ,δ_2–$^2H_7$;α–^{15}N]-leucine, "δ_1-Leu", or L-[δ_2–^{13}C;α,β,γ,δ_1–$^2H_7$;α–^{15}N]-leucine, "δ_2-Leu", : L-[γ_1–^{13}C;α,β,γ_2–$^2H_5$;α–^{15}N]-valine, "γ_1-Val", or L-[γ_2–^{13}C;α,β,γ_1–$^2H_5$;α–^{15}N]-valine, "γ_2-Val". We successfully prepared 82 kDa MSG samples selectively labeled with either δ_1-Leu, δ_2-Leu, γ_1-Val, or γ_2-Val, respectively, in otherwise fully deuterated and uniformly ^{15}N labeled backgrounds. By choosing any of the two pairs from the stereo-specifically $^{13}CH_3$-labeled leucines and valines, we could prepare combinatorially labeled MSG samples with either the δ_2-Leu + γ_2-Val, δ_2-Leu + γ_1-Val, δ_1-Leu + γ_2-Val, or δ_1-Leu + γ_1-Val pair, in otherwise fully deuterated backgrounds. The incorporation efficiencies for the Leu and Val residues in the residue-selective, or combinatorially Leu and Val labeled MSG samples were considerably higher than 90 % (for Leu) and 80 % (for Val), even with relatively small amounts of labeled leucines ($\sim$20 mg/L) and valines ($\sim$100 mg/L) added to the medium. The differential, stereo-selective labeling method for Leu and Val residues using the cellular expression system is widely applicable for preparing isotope-labeled proteins for NMR studies, and may also be compatible with the precursor method. Note that one may also incorporate labeled amino acids other than leucine and valine, by modifying the *E. coli* cellular expression protocol described in this article. The production of collectively isotope-labeled proteins with selected sets of amino acids bearing sophisticated labeling patterns, such as SAIL amino acids, without using cell-free protein expression, will accelerate the development of practical methodologies for studying the structures and dynamics of large proteins and protein complexes. The stereo-specifically $^{13}CH_3$ labeled leucines and valines used in this study will soon be commercially available from SAIL Technologies Inc. (http://www.sail-technologies.com/).

Acknowledgments This work was supported by Platform for Drug Discovery, Informatics, and Structural Life Science (MEXT), and also in part by the Targeted Protein Research Program (MEXT). Y. M. was supported by JSPS KAKENHI Grants-in-Aid for Young Scientists (B) (23770111 and 25840021), and M. T. was supported by a JSPS KAKENHI Grant-in-Aid for Young Scientists (B) (23770109), a Grant-in-Aid for Scientific Research (C) (25440018), and a grant from the Kurata Memorial Hitachi Science and Technology Foundation. We would like to thank Dr. Tugarinov, of the University of Maryland, for providing the MSG gene and his advice regarding MSG protein production. We also thank Dr. Jérôme Boisbouvier, of Institute de Biologie Structurale Jean-Pierre Ebel, for valuable discussions.

References

Amero C, Schanda P, Durá MA, Ayala I, Marion D, Franzetti B, Brutscher B, Boisbouvier J (2009) Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies. J Am Chem Soc 131:3448–3449

Amero C, Asunción DM, Noirclerc-Savoye M, Perollier A, Gallet B, Plevin MJ, Vernet T, Franzetti B, Boisbouvier J (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. J Biomol NMR 50:229–236

Ardá A, Jiménez C, Rodríguez J (2004) A study of polychlorinated leucine derivatives: synthesis of (2*S*,4*S*)-5,5-dichloroleucine. Tetrahedron Lett 45:3241–3243

August RA, Khan JA, Moody CM, Young DW (1992) Stereospecific synthesis of (2*S*,4*R*)-[5,5,5-$^2H_3$]-leucine. Tetrahedron Lett 33:4617–4620

Ayala I, Sounier R, Usě N, Gans P, Boisbouvier J (2009) An efficient protocol for the complete incorporation of methyl-protonated alanine in perdeuterated protein. J Biomol NMR 43:111–119

Charrier J, Hitchcock PB, Young DW (2004) Synthesis of (2*S*,3*R*)-[3′,3′,3′-$^2H_3$]-valine and [2*S*,3*S*]-4-fluorovaline. Org Biomol Chem 2:1310–1314

Dawadi PBS, Lugtenburg J (2013) Access to any site directed stable isotope (^{2}H, ^{13}C, ^{15}N, ^{17}O and ^{18}O) in genetically encoded amino acids. Molecules 18:482–519

Fischer M, Kloiber K, Häusler J, Ledolter K, Konrat R, Schmid W (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

Gans P, Hamelin O, Sounier R, Ayala I, Dura MA, Amero CD, Noirclerc-Savoye M, Franzetti B, Plevin MJ, Boisbouvier J (2010) Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. Angew Chem 122:2002–2006

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

Gelis I, Bonvin AMJJ, Keramisanou D, Koukaki M, Gouridis G, Karamanou S, Economou A, Kalodimos CG (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. Cell 131:756–769

Gill ML, Palmer AG (2011) Multiplet-filtered and gradient-selected zero-quantum TROSY experiments for 13C1H3 methyl groups in proteins. J Biomol NMR 51:245–251

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

Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE (1999) A robust and cost-effective method for the production of Val, Leu, Ile(δ1) methyl-protonated ^{15}N-, ^{13}C-, ^{2}H-labeled proteins. J Biomol NMR 13:369–374

Guo C, Tugarinov V (2010) Selective ^{1}H-^{13}C NMR spectroscopy of methyl groups in residually protonated samples of large proteins. J Biomol NMR 46:127–133

Howard BR, Endrizzi JA, Remington SJ (2000) Crystal Structure of *Escherichia coli* Malate Synthase G complexed with magnesium and glyoxylate at 2.0 Å resolution: mechanistic implications. Biochemistry 39:3156–3168

Hu H, Namanja AT, Wong S, Chen Y (2012) Selective editing of Val and Leu methyl groups in high molecular weight protein NMR. J Biomol NMR 53:113–124

Humphrey JM, Hart JA, Chamberlin AR (1995) Efficient syntheses of diastereoselectively labeled (3*S*)-[4-^{13}C]-L-valine, and regioselectively labeled [3-$^{13}CH_3$]-L-isoleucine hydrochlorides. Bioorg Med Chem Lett 5:1315–1320

Isaacson RL, Simpson PJ, Liu M, Cota E, Zhang X, Freemont P, Matthews S (2007) A new labeling method for methyl transverse relaxation-optimized spectroscopy NMR spectra of alanine residues. J Am Chem Soc 129:15428–15429

Janin J, Miller S, Chothia C (1988) Surface, subunit interfaces and interior of oligomeric proteins. J Mol Biol 204:155–164

Kainosho M, Güntert P (2009) SAIL–stereo-array isotope labeling. Q Rev Biophys 42:247–300

Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Ono AM, Güntert P (2006) Optimal isotope labelling for NMR protein structure determinations. Nature 440:52–57

Kelly NM, Sutherland A, Willis CL (1997) Syntheses of amino acids incorporating stable isotopes. Nat Prod Rep 14:205–219

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

Mas G, Crublet E, Hamelin O, Gans P, Boisbouvier J (2013) Specific labeling and assignment strategies of valine methyl groups for the NMR studies of high molecular weight proteins. J Biomol NMR. doi:10.1007/s10858-013-9785-z

Metzler WJ, Wittekind M, Goldfarb V, Mueller L, Farmer BT (1996) Incorporation of $^{1}H/^{13}C/^{15}N$-{Ile, Leu, Val} into a perdeuterated, ^{15}N-labeled protein: potential in structure determination of large proteins by NMR. J Am Chem Soc 118:6800–6801

Miyanoiri Y, Takeda M, Jee JG, Ono AM, Okuma K, Terauchi T, Kainosho M (2011) Alternative SAIL-Trp for robust aromatic signal assignment and determination of the χ_2 conformation by intra-residue NOEs. J Biomol NMR 51:425–435

Miyanoiri Y, Takeda M, Kainosho M (2012) Stereo-array isotope labeling method for studying protein structure and dynamics. Adv Exp Med Biol 992:83–93

Molina I, Pellicer MT, Badia J, Aguilar J, Baldoma L (1994) Molecular characterization of *Escherichia coli* malate synthase G: Differentiation with the malate synthase A isoenzyme. Eur J Biochem 224:541–548

Neri D, Szyperski T, Otting G, Senn H, Wüthrich K (1989) Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional carbon-13 labeling. Biochemistry 28:7510–7516

Oba M, Ueno R, Fukuoka M, Kainosho M, Nishiyama K (1995) Synthesis of L-threo- and L-erythro-[1-^{13}C,2,3-$^{2}H_2$] amino-acids: Novel probes for conformational analysis of peptide side chains. J Chem Soc Perkin Trans 1:1603–1609

Oba M, Terauchi T, Miyakawa AA, Kamo H, Nishiyama K (1998) Stereoselective deuterium-labelling of diastereotopic methyl and methylene protons of L-leucine. Tetrahedron Lett 39:1595–1598

Otten R, Chu B, Krewulak KD, Vogel HJ, Mulder FAA (2010) Comprehensive and cost-effective NMR spectroscopy of methyl groups in large proteins. J Am Chem Soc 132:2952–2960

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

Rosen MK, Gardner KH, Willis RC, Parris WE, Pawson T, Kay LE (1996) Selective methyl group protonation of perdeuterated proteins. J Mol Biol 263:627–636

Shekhtman A, Ghose R, Goger M, Cowburn D (2002) NMR structure determination and investigation using a reduced proton (REDPRO) labeling strategy for proteins. FEBS Lett 524:177–182

Sheppard D, Guo C, Tugarinov V (2009a) 4D, pp. 1H–13C NMR spectroscopy for assignments of alanine methyls in large and complex protein structures. J Am Chem Soc 131:1364–1365

Sheppard D, Guo C, Tugarinov V (2009b) Methyl-detected 'out-and-back' NMR experiments for simultaneous assignments of Alaβ and Ileγ2 methyl groups in large proteins. J Biomol NMR 43:229–238

Skrynnikov NR, Mulder FAA, Hon B, Dahlquist FW, Kay LE (2001) Probing slow time scale dynamics at methyl-containing side chains in proteins by relaxation dispersion NMR measurements: application to methionine residues in a cavity mutant of T4 lysozyme. J Am Chem Soc 123:4556–4566

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

Takeda M, Ikeya T, Güntert P, Kainosho M (2007) Automated structure determination of proteins with the SAIL-FLYA NMR method. Nat Protoc 2:2896–2902

Takeda M, Hallenga K, Shigezane M, Waelchli M, Löhr F, Markley JL, Kainosho M (2011) Construction and performance of an NMR tube with a sample cavity formed within magnetic susceptibility matched glass. J Magn Reson 209:167–173

Takeda M, Terauchi T, Kainosho M (2012) Conformational analysis by quantitative NOE measurements of the β-proton pairs across individual disulfide bonds in proteins. J Biomol NMR 52:127–139

Torizawa T, Shimizu M, Taoka M, Miyano H, Kainosho M (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. J Biomol NMR 30:311–325

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

Tugarinov V, Kay LE (2004a) An isotope labeling strategy for methyl TROSY spectroscopy. J Biomol NMR 28:165–172

Tugarinov V, Kay LE (2004b) Stereospecific NMR assignments of prochiral methyls, rotameric states and dynamics of valine residues in malate synthase G. J Am Chem Soc 126:9827–9836

Tugarinov V, Kay LE (2005) Methyl groups as probes of structure and dynamics in NMR studies of high-molecular-weight proteins. ChemBioChem 6:1567–1577

Tugarinov V, Muhandiram R, Ayed A, Kay LE (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: Chemical shift assignments and secondary structure of malate synthase G. J Am Chem Soc 124:10025–10035

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

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

Velyvis A, Kay LE (2013) Measurement of active site ionization equilibria in the 670 kDa Proteasome core particle using methyl-TROSY NMR. J Am Chem Soc 135:9259–9262

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

Xu Y, Matthews S (2013) TROSY NMR spectroscopy of large soluble proteins. Top Curr Chem 335:97–119