Isotope Labeling of Alanine Methyl Groups on a Deuterated Background for NMR Studies of High-Molecular-Weight Proteins

Chenyun Guo, Raquel Godoy-Ruiz and Vitali Tugarinov
Department of Chemistry and Biochemistry
University of Maryland, College Park, MD 20742 USA
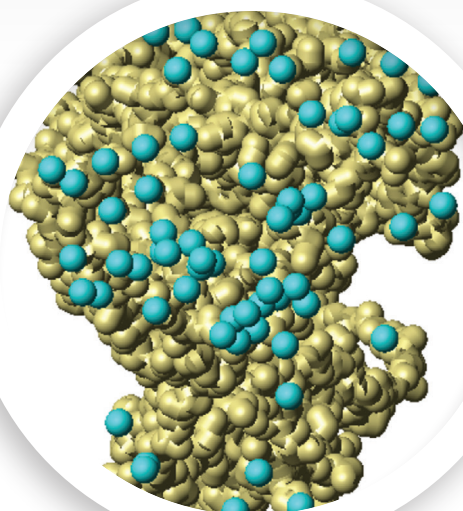
The state-of-the-art isotope-labeling schemes commonly employed for NMR investigations of high-molecular-weight proteins utilize selective incorporation of protons and ^{13}C isotopes into methyl groups of Ile$^{\delta 1}$, Leu$^{\delta}$ and Val$^{\gamma}$ side chains in a highly deuterated environment (commonly referred to as "ILV labeling") providing a large number of high-quality probes for NMR studies of protein structure and dynamics. Robust ILV labeling methods[1,2] and strategic location of ILV side-chains in the hydrophobic cores of protein structures[3,4] have turned ILV labeling into an indispensable tool for NMR studies of large proteins and macromolecular assemblies.[5]

More than a decade ago, Gardner & Kay pioneered a technique for selective incorporation of protonated Ile$^{\delta 1}$ methyl positions into protein molecules.[6] Subsequently, selective methyl labeling/ protonation methods have been refined and extended to Leu$^{\delta}$ and Val$^{\gamma}$ methyl sites by the Kay group.[7,8] The introduction of methyl transverse relaxation optimized spectroscopy[9] (methyl-TROSY) has stimulated the development of improved labeling schemes.[10,11] Selective labeling of ILV methyl sites in large proteins on a deuterated background in synergy with methyl-TROSY techniques have had a significant impact on NMR studies of structure and dynamics of large protein assemblies up to ~800 kDa in molecular weight.[12-15] Nevertheless, the availability of only three ILV probes oftentimes presents a serious limitation for structural and dynamics studies of large proteins. Recent advances in methyl isotope labeling have focused on the extension of the ILV labeling methodology to methyl positions of alanines (Ala$^{\beta}$).

Methyl groups of alanines serve as an attractive extension of the ILV labeling methodology. Ala methyls are located in close proximity to the backbone; their flexibility is therefore reduced compared to other methyl-bearing side chains in proteins. Ala is one of the most frequently encountered residues in protein hydrophobic cores and at molecular interfaces.[16] Recently, several reports have focused upon extending the ILV methodology with (selective) labeling of Ala$^{\beta}$ methyl positions in large proteins. Isaacson et al. have introduced a labeling strategy for selective incorporation of ^{13}CH$_3$ groups into alanine residues on a deuterated background.[17] Specific ^{13}C labeling and protonation of Ala$^{\beta}$ methyls to a level of 95% with minimal background labeling (<1%) in minimal D$_2$O-based bacterial medium
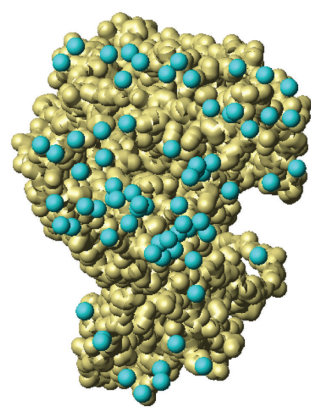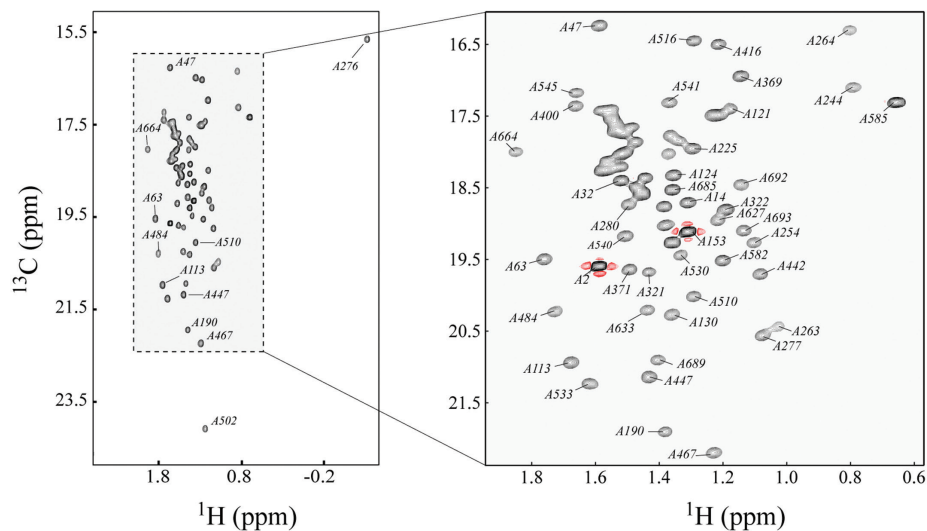
Figure 1. Left: Schematic representation of the structure of malate synthase G (82-kDa; 73 alanines) with positions of Ala[β] carbon atoms shown with cyan-colored spheres. Center: Methyl ^{1}H-^{13}C HMQC correlation map recorded on a 0.9 mM {U-^{2}H; Ala[β]-[^{13}CH$_3$]}-labeled MSG (37 °C; 600 MHz). Right: The "zoomed" part of the correlation map highlighted in the center. Selected assignments of separated Ala[β] cross-peaks are indicated.[22] Negative contours are shown in red.

supplemented with large amounts of selectively ^{13}C-labeled α-deuterated alanine and co-addition of three deuterated compounds: (i) α-ketoisovalerate-D$_7$, (ii) succinate-D$_4$ and (iii) L-isoleucine-D$_{10}$ has been reported by Boisbouvier and coworkers.[18] This labeling protocol has been closely followed for production of {^{2}H; Ala[β]-[^{13}CH$_3$]}-labeled malate synthase G[19,20] (MSG) – an 82-kDa enzyme containing 73 methyl groups. Alanine is the most abundant residue in MSG comprising 10.1% of the total amino-acid content. Figure 1 shows the methyl-TROSY ^{1}H-^{13}C correlation map of MSG prepared using the protocol described in more detail in Appendix 1.

To maximize the number of available methyl probes, it is clearly advantageous to combine selective labeling of Ala positions with ILV methyl labeling.[21] The use of selectively ^{13}CH$_3$-labeled Ala[β] ({2-D, 3-^{13}C}-L-alanine) in combination with (i) selectively ^{13}CH$_3$-labeled α-ketobutyrate for labeling of Ile[δ1] poistions, (ii) ^{13}CH$_3$/^{12}CD$_3$-labeled α-ketoisovalerate for labeling of Val[γ] and Leu[δ] sites, and (iii) deuterated succinate, achieves exactly this purpose (see Appendix 1 for the details about the used compounds). Figure 2 shows the ^{1}H-^{13}C methyl-TROSY correlation map recorded on the {Ala[β]-[^{13}CH$_3$]; Ile[δ1]-[^{13}CH$_3$]; Leu,Val-[^{13}CH$_3$/^{12}CD$_3$]}-labeled MSG. Although the overlap of resonances within the Ala[β] group is very substantial, in MSG Ala[β] signals do not overlap with either Val[γ] or Leu[δ] methyl groups.

In full agreement with previous observations,[18] no signs of scrambling of isotope labels from alanine to other amino-acids have been detected in either sample. In fact, a comparison of signal intensities of ILV methyls in the {U-^{2}H; Ala[β]-[^{13}CH$_3$]; Ile[δ1]-[^{13}CH$_3$]; Leu,Val-[^{13}CH$_3$/^{12}CD$_3$]}-labeled sample with those obtained in {U-^{2}H; Ile[δ1]-[^{13}CH$_3$]; Leu,Val-[^{13}CH$_3$/^{12}CD$_3$]}-labeled MSG (without

Ala[β] labeling) indicated that the levels of isotope incorporation into ILV methyl positions are not compromised to any significant extent by additions of large amounts of labeled Ala to the medium. This is the direct consequence of the fact that the biosynthetic pathway leading to incorporation of labels into Ile[δ1] positions is "short-circuited" by addition of the suitably labeled α-ketobutyrate to the medium, while the pathway leading to labeling of Val[γ] and Leu[δ] sites is "short-circuited" by addition of the α-keto-isovalerate precursor. It might have been expected that the addition of α-ketobutyrate to the medium instead of deuterated isoleucine in order to ensure ^{13}CH$_3$ labeling at Ile[δ1] methyl positions, would result in partial ^{13}CH$_3$ labeling at Ile[γ2] methyl sites (arising from alanine-derived [3-^{13}CH$_3$]-pyruvate entering into the biosynthetic cycle of Ile[δ] instead of the completely deuterated pyruvate from [U-^{2}H]-glucose). Nevertheless, Ile[γ2] ^{1}H-^{13}C correlations have not been detected in the ^{1}H-^{13}C methyl-TROSY spectra recorded on the {Ala[β]-[^{13}CH$_3$]; Ile[δ1]-[^{13}CH$_3$]; Leu,Val-[^{13}CH$_3$/^{12}CD$_3$]}-MSG indicating that during protein production the molecules of pyruvate originate predominantly from [U-^{2}H]-glucose.

Proximity of alanine side chains to the protein backbone and their high degree of order turn Ala[β] methyl groups into excellent NMR probes for a number of applications: (i) measurements of ^{1}H-^{13}C residual dipolar couplings (RDCs) in Ala[β] methyls, (ii) characterization of fast (pico-to-nanosecond) and slow (μs-to-millisecond) dynamics at functionally important sites of enzymes, and (iii) methyl-TROSY NOE spectroscopy that can be performed on {U-^{2}H; Ala[β]–[^{13}CH$_3$]; Ile[δ1]-[^{13}CH$_3$]; Leu,Val-[^{13}CH$_3$/^{12}CD$_3$]}-labeled samples increasing the number of methyl probes for derivation of distance restraints compared to conventional ILV labeling methodology.
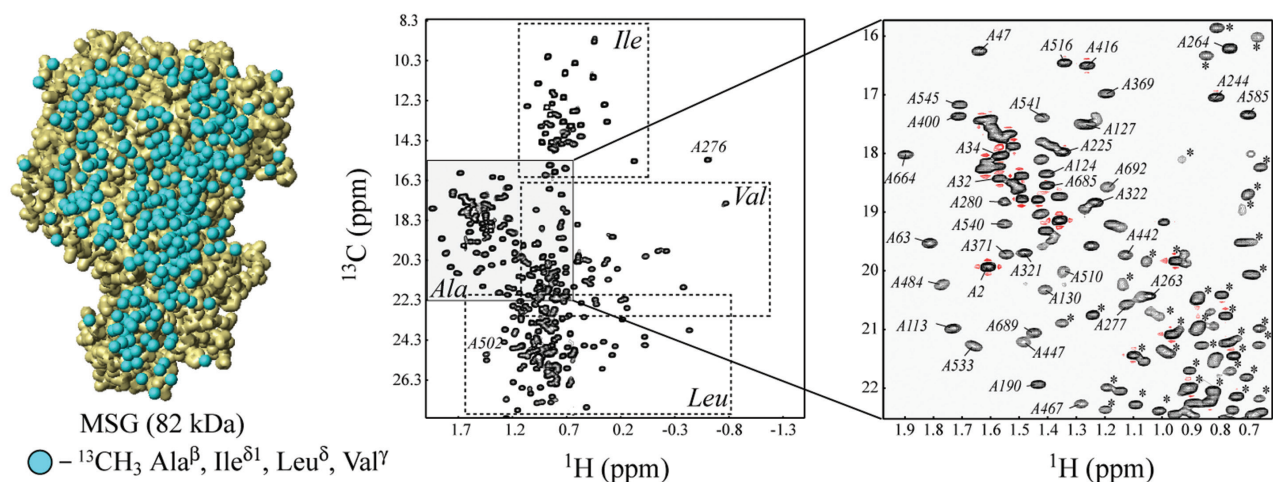
Figure 2. Left: Schematic representation of the structure of MSG with methyl carbons of Ala[β], Ile[δ1], Leu[δ] and Val[γ] positions shown with cyan-colored spheres. Center: Methyl ^{1}H-^{13}C HMQC correlation map recorded on a 0.75 mM {Ala[β]-[^{13}CH$_3$]; Ile[δ1]-[^{13}CH$_3$]; Leu,Val-[^{13}CH$_3$/^{12}CD$_3$]}-labeled MSG (37°C; 600 MHz). The regions of the map approximately corresponding to Ile[δ1], Val[γ] and Leu[δ] methyl positions are enclosed in dashed rectangles. The region enclosed in a solid rectangle and highlighted corresponds to Ala[β] methyl correlations. Right: The Ala[β] part of the map "zoomed" from the region highlighted in the center. Cross-peaks arising from Val[γ] and Leu[δ] correlations are marked with asterisks. Ala[β] assignments are indicated for selected methyls.[22]

Appendix 1. Selective labeling of Ala[β] positions with ^{13}CH$_3$ groups (the {U-[^{2}H]; Ala[β]-[^{13}CH$_3$]}-labeled sample) has been achieved following the protocol of Ayala *et al.*[18] using [U-^{2}H]-D-glucose as the main carbon source in *E. coli* medium and addition of (i) 800 mg of {2-D, 3-^{13}C}-L-alanine, (ii) 2.5 g of succinate-D$_4$, (iii) 200 mg for α-keto-isovalerate-D$_7$ and (iv) 60 mg of L-isoleucine-D$_{10}$ to 1 L of D$_2$O-based M9 medium 1 hour prior to induction of protein overexpression with 1 mM IPTG. Selective ^{13}CH$_3$ labeling of all ILV positions together with Ala[β] methyls ({U-[^{2}H]; Ala[β]-[^{13}CH$_3$]; Ile[δ1]-[^{13}CH$_3$]; Leu,Val-[^{13}CH$_3$/^{12}CD$_3$]}-labeled sample) has been achieved using the same carbon sources as above by addition of (i) 800 mg of {2-D,3-^{13}C}-L-alanine; (ii) 2.5g of succinate-D$_4$; (iii) 120 mg of α-ketoisovaleric acid, sodium salt (3-methyl-^{13}C; 3,4,4,4-D$_4$) and (iv) 60 mg of α-Ketobutyric acid, sodium salt (3-methyl-^{13}C; 3,3-D$_2$) to 1 L of the medium 1 hour prior to induction. {Ala[β]-[^{13}CH$_3$]}- and {Ala[β]-[^{13}CH$_3$]; Ile[δ1]-[^{13}CH$_3$]; Leu,Val-[^{13}CH$_3$/^{12}CD$_3$]}-labeled samples of MSG were 0.9 mM and 0.75 mM in protein concentration, respectively, in 99% D$_2$O and contained 25 mM sodium phosphate buffer (pH 7.1, uncorrected), 20 mM MgCl$_2$, 5 mM DTT and 0.05% NaN$_3$.

CIL Products

IN VIVO ILVA METHYL LABELING KIT
CIL Catalog # CDLM-8806-KIT

This kit contains four separate vials of the following compounds in the amounts indicated:

Compound Name	Catalog No.	Amount
Succinate (D_4)	DLM-584	2.5 g
α-Ketoisovaleric acid, sodium salt (3-methyl-^{13}C; 3,4,4,4-D_4)	CDLM-7317	0.12 g
α-Ketobutyric acid, sodium salt (3-methyl-^{13}C; 3,3-D_2)	CDLM-7318	0.06 g
Alanine (2-D, 3-^{13}C)	CDLM-8649	0.8 g

The materials in this kit are to be used in conjunction with 1 L of deuterated minimal media. 1 L of minimal media that is suitable for use with this product typically contains 2 g of glucose-D_7, 1 g of ammonium salt (either ^{15}N labeled or unlabeled, depending on preference of the user), and 11.3 g of M9 salts in D_2O.

IN VIVO ALANINE METHYL LABELING KIT
CIL Catalog # CDLM-8805-KIT

This kit contains four separate vials of the following compounds in the amounts indicated:

Compound Name	Catalog No.	Amount
Succinate (U-D_4)	DLM-584	2.5 g
α-Keto Isovalerate (U-D_7)	DLM-4646	0.2 g
Isoleucine (U-D_{10})	DLM-141	0.06 g
Alanine (2-D, 3-^{13}C)	CDLM-8649	0.8 g

The materials in this kit are to be used in conjunction with 1 L of deuterated minimal media.[1,2] 1 L of minimal media that is suitable for use with this product typically contains 2 g of glucose-D_7, 1 g of ammonium salt (either ^{15}N labeled or unlabeled, depending on preference of the user), and 11.3 g of M9 salts in D_2O.

[1] Ayala, I.; Sounier, R.; Use, N.; Gans, P.; Boisbouvier, J. **2009**. *J Biomol NMR, 43,* 111-119.

[2] Isaacson, R.L.; Simpson, P.J.; Liu, M.; Cota, E.; Zhange, X.; Freemont, P.; Matthews, S. **2009**. *J Am Chem Soc, 129,* 15428-15429.

Don't see your desired labeling pattern? Want pre-weighed labeled components for your media preparation? We can usually package our reagents to your specifications at no extra charge – please see your local sales representative for a quote.

References

1. Tugarinov, V.; Kanelis, V.; Kay, L.E. **2006**. *Nat Protocols, 1,* 749.

2. Sprangers, R.; Velyvis, A.; Kay, L.E. **2007**. *Nat Methods, 4,* 697.

3. Gardner, K.H.; Rosen, M.K.; Kay, L.E. **1997**. *Biochemistry, 36,* 1389.

4. Rosen, M.K.; Gardner, K.H.; Willis, R.C.; Parris, J.E.; Pawson, T.; Kay, L.E. **1996**. *J Mol Biol, 263,* 627.

5. Tugarinov, V.; Hwang, P.M.; Kay, L.E. **2004**. *Annu Rev Biochem, 73,* 107.

6. Gardner, K.H.; Kay, L.E. **1997**. *J Am Chem Soc, 119,* 7599.

7. Gardner, K.H.; Zhang, X.; Gehring, K.; Kay, L.E. **1998**. *J Am Chem Soc, 120,* 11738.

8. Goto, N.K.; Gardner, K.H.; Mueller, G.A.; Willis, R.C.; Kay, L.E. **1999**. *J Biomol NMR, 13,* 369.

9. Tugarinov, V.; Hwang, P.M.; Ollerenshaw, J.E.; Kay, L.E. **2003**. *J Am Chem Soc, 125,* 10420.

10. Tugarinov, V.; Kay, L.E. **2004**. *J Biomol NMR, 28,* 165.

11. Gans, P.; Hamelin, O.; Sounier, R.; Ayala, I.; Durá, M.A.; Amero, C.D.; Noirclerc-Savoye, M.; Franzetti, B.; Plevin, M.J.; Boisbouvier, J. Angew. **2010**. *Chem Int Ed Engl, 49,* 1958.

12. Tugarinov, V.; Choy, W.Y.; Orekhov, V.Y.; Kay, L.E. **2005**. *Proc Natl Acad Sci USA, 102,* 622.

13. Sprangers, R.; Kay, L.E. **2007**. *Nature, 445,* 618.

14. Religa, T.L.; Sprangers, R.; Kay, L.E. **2010**. *Science, 328,* 98.

15. Velyvis, A.; Yang, Y.R.; Schachman, H.K.; Kay, L.E. **2007**. *Proc Natl Acad Sci USA, 104,* 8815.

16. Fernandez, A.; Scott, L.R.; Scheraga, H.A. **2003**. *J Phys Chem, 107,* 9929.

17. Isaacson, R.L.; Simpson, P.J.; Liu, M.; Cota, E.; Zhang, X.; Freemont, P.; Matthews, S. **2007**. *J Am Chem Soc, 129,* 15428.

18. Ayala, I.; Sounier, R.; Usé, N.; Gans, P.; Boisbouvier, J. **2009**. *J. Biomol. NMR, 43,* 111.

19. Howard, B.R.; Endrizzi, J.A.; Remington, S. **2000**. *J. Biochemistry, 39,* 3156.

20. Tugarinov, V.; Muhandiram, R.; Ayed, A.; Kay, L.E. **2002**. *J Am Chem Soc, 124,* 10025.

21. Godoy-Ruiz, R.; Guo, C.; Tugarinov V. **2010**. *J Am Chem Soc, 132*(51) 18340.

22. Sheppard, D.; Guo, C.; Tugarinov V. **2009**. *J Am Chem Soc, 131,* 1364.