J|A|C|S COMMUNICATIONS

Published on Web 01/08/2009

4D $^{1}H-^{13}C$ NMR Spectroscopy for Assignments of Alanine Methyls in Large and Complex Protein Structures

Devon Sheppard, Chenyun Guo, and Vitali Tugarinov*

Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742

Received October 18, 2008; E-mail: vitali@umd.edu

Selective $^{13}CH_3$ labeling of Ile(δ1), Leuδ, and Valγ (ILV) positions in large proteins on a deuterated background[1,2] coupled with Methyl-TROSY[3] have significantly extended the range of applicability of current protein NMR methodology and facilitated NMR studies of structure and dynamics in very large protein systems.[4−7] Clearly, the availability of only these three probes represents a serious limitation for structural and dynamics studies of large proteins. In search of additional NMR probes we have targeted methyl groups of alanines as potentially an attractive extension of the ILV labeling methodology. Ala methyls are located in close proximity to the backbone; their structural "degrees of freedom" are, therefore, reduced compared to other methyl sites in proteins. Alanine is one of the most abundant amino acids found in proteins;[8] it is one of the most frequently encountered residues both in protein hydrophobic cores and at molecular interfaces.[9] Resonance assignments of Ala methyls in large and complex proteins pose significant difficulties even if backbone $^{13}C/^{15}N$ chemical shifts are available from prior studies. With the importance of alanines as structural probes in mind, we have developed an NMR experiment for assignments of Ala $^{13}CH_3$ methyls in large proteins and applied it to Malate Synthase G (MSG), a 723-residue monomeric enzyme extensively studied by NMR earlier.[10]

Recently, an isotope labeling strategy has been introduced for selective labeling of Ala methyls on a deuterated background.[11] Since MSG can be obtained in sufficient quantities even when pyruvate is used as the main carbon source, we have chosen to use {U$-^{13}$C}-Pyruvate as the carbon source in a D_2O-based minimal medium to generate protonated Ala methyls in MSG.[12,13] The clear disadvantage of this labeling strategy is that $^{13}CH_3$ methyl isotopomers are generated only at the maxumim level of ~25% (of the total isotopomer content in MSG). $^{13}CHD_2$ was, however, the only other methyl isotopomer observable in the spectra of {U$-^{13}$C}-Pyruvate-derived MSG. In addition to Ala methyls, the {U$-^{13}$C}-Pyruvate-derived samples of MSG contained (partially) protonated Ileγ2, Valγ, Leuδ methyl sites and C$^{\gamma}$(C$^{\beta}$) positions in Arg, Pro, Glx(Asx, Ser).[12,13]

Alanine is the most abundant residue in MSG. With 73 alanines (10.1% of amino-acid content), MSG is an especially challenging system for assignments of Ala methyls. Ala methyl resonances tend to be very ill-dispersed in proteins, with 88% of Ala methyls in MSG clustered between 1.0 and 1.7 ppm (17 and 21 ppm) in the ^{1}H(^{13}C) dimension. As Ala ^{13}C chemical shifts of a deuterated form of MSG are known from previous studies,[10] the assignment problem seemingly boils down to "positioning" of Ala methyls on a 2D $^{1}H-^{13}C$ correlation map according to their known $^{13}C^{\beta}$ shifts. However, even this task can not be accomplished without additional experiments due to degeneracy of many Ala $^{13}C^{\beta}$ shifts. The assignment problem is aggravated by the fact that $^{13}C^{\beta}$ chemical shifts in Ala $^{13}CH_3$ groups differ from those of $^{13}CD_3$ methyls by three times the value of a one-bond deuterium isotope shift[14] ($^{1}\Delta_{D\rightarrow H} \approx 0.3$ ppm) that is uniform, in our experience, only to within $\pm$0.1 ppm. Therefore, protonated $^{13}C^{\beta}$ shifts in alanines of MSG are known only approximately.

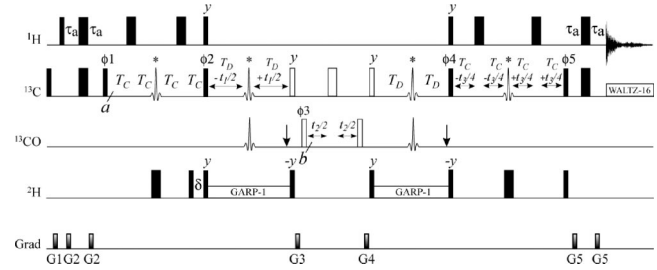

Figure 1. 4D ALA-HMCBCACO methyl-detected "out-and-back" pulse-scheme for assignment of Ala methyls in large proteins. All narrow(wide) rectangular pulses are applied with flip angles of 90°(180°) along the *x*-axis unless indicated otherwise. All the pulses shown with black rectangles are applied with the highest possible power. ^{13}C WALTZ-16[16] decoupling is achieved using a 2 kHz field, while ^{2}H GARP-1[17] decoupling uses a 0.9 kHz field. The ^{1}H(^{2}H) carrier is positioned at 4.7(2.5) ppm. The ^{13}C carrier is placed at 18 ppm (the center of the Ala methyl region), switched to 36 ppm before the ^{13}C pulse with phase ϕ1, to 177 ppm before the pulse with phase ϕ3, and then returned to 36 ppm prior to the second open 90° ^{13}C pulse with phase *y*, and back to 18 ppm after the pulse with phase ϕ5. The 90°(180°) ^{13}C and ^{13}CO pulses shown with open rectangles are applied with a field strength of $\Delta/\sqrt{15}(\Delta/\sqrt{3})$ where Δ is the difference (in Hz) between $^{13}C^{\alpha}$(53 ppm) and ^{13}CO(177 ppm) chemical shifts.[18] The ^{13}C pulses shown with open rectangles are applied at 53 ppm by phase-modulation of the carrier.[19] The ^{13}C shaped pulses (marked with *) are 400 μs(600 MHz) RE-BURP[20] pulses. It is important to adjust the phases of these RE-BURP pulses to maximize sensitivity. The ^{13}CO shaped pulses are 400 μs RE-BURP pulses applied at 177 ppm by phase modulation of the carrier. Vertical arrows at the end of the t_1 and $2T_D$ periods indicate the position of the ^{13}CO Bloch-Siegert compensation pulse.[18] Delays are: $\tau_a = 2.0$ ms; $T_C = 3.5$ ms; $T_D = 4.75$ ms; $\delta = 1.4$ ms. The phase-cycle is: $\phi1 = x, -x$; $\phi2 = 2(y), 2(-y)$; $\phi3 = 4(x), 4(-x)$; $\phi4 = 4(y), 4(-y)$; $\phi5 = x$; rec $= 2(x,-x), 2(-x,x)$. Quadrature in $F_1(F_2)$ and F_3 is achieved by States-TPPI[21] of $\phi1,\phi2(\phi3)$ and $\phi5$, respectively. Durations and strengths of the pulsed-field gradients in units of (ms;G/cm) are: G1 = (1.0; 5), G2 = (0.3; 8), G3 = (0.8; 10), G4 = (1.2; 12), G5 = (0.3; 15).

Figure 1 shows the pulse scheme for a 4D ALA-HMCBCACO experiment used for assignment of Ala methyls in MSG. The pulse scheme provides intraresidual correlations of the form ($\Omega_{C\alpha}$, Ω_{CO}, $\Omega_{C\beta}$, $\Omega_{H\beta}$) in alanines and is based upon the methyl-detected "out-and-back" methodology developed earlier for ILV assignments.[15] The flow of magnetization from point *a* to point *b* can be represented as

$$2H_x^m C_y^{\beta} \xrightarrow{4T_C,\ 90_y,\ 90_{\phi 2}} 4H_z^m C_z^{\beta} C_x^{\alpha} \xrightarrow{2T_D(t_1)} 4H_z^m C_z^{\alpha} C'_z(t_1) \xrightarrow{90_y,\ 90_{\phi 3}} 4H_z^m C_z^{\alpha} C'_y(t_2) \quad (1)$$

where C_r^i and $H_r^m = H_r^1 + H_r^2 + H_r^3$ are product operators for carbon *i* and the three methyl proton nuclei, respectively, $r = x, y, z$, and we concentrate only on the part of magnetization that leads to an

10.1021/ja808202q CCC: $40.75

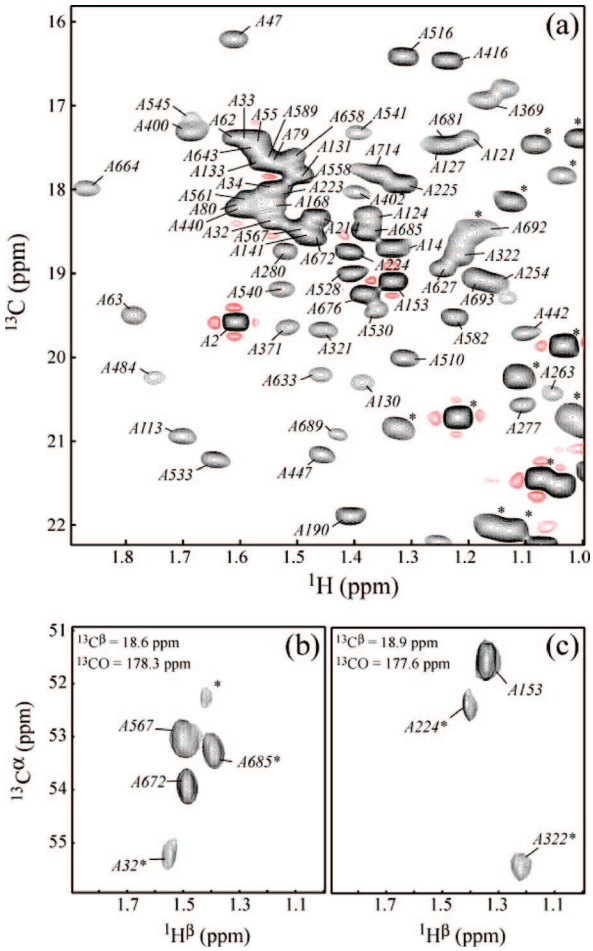

Figure 2. (a) Alanine region of a 2D methyl ^{1}H−^{13}C CT-HMQC spectrum recorded on a 0.9 mM {U−^{13}C}-Pyruvate-derived sample of MSG (99.9% D_2O; 600 MHz; 37 °C; pH = 7.1). The spectrum was "filtered" for $^{13}CH_3$ (against $^{13}CHD_2$) methyl groups.[12] Assignments of Ala methyls are indicated with residue numbers. The peaks belonging to Valγ and Ileγ2 methyl signals are marked with *. (b−c) Selected 2D $^{1}H^{\beta}$−$^{13}C^{\alpha}$ slices from the 4D ALA-HMCBCACO data set recorded on the same sample as in (a) using the pulse scheme of Figure 1. The planes are drawn at $^{13}C^{\beta}$/^{13}CO chemical shifts indicated at the top left corner of each panel. Methyl cross-peaks are marked with residue numbers, while the peaks whose maxima lie outside of the plotted 2D slices are marked with *.

observable signal at the end of the scheme. The pulse scheme utilizes the Methyl-TROSY[3] principle where possible by keeping the methyl ($^{1}H^{\beta}$−$^{13}C^{\beta}$) magnetization in a multiple-quantum state before the $^{13}C_{\phi 2}/H_y$ pair of pulses and after the $^{13}C_{\phi 4}/H_y$ pair of pulses (Figure 1). To eliminate $^{13}CHD_2$ methyl isotopomer signals from the spectra, the 180° ^{2}H pulses in the middle of $4T_C$ and t_3 periods ensure evolution of the $^{1}J_{C\text{-}D}$ coupling; subsequent application of ^{2}H 90° pulses at time points approximately equal to 0.25/$^{1}J_{C\text{-}D}$ eliminates most of the deuterium-containing magnetization terms in $^{13}CHD_2$ groups.[12]

Chemical shift degeneracy in $^{13}C^{\alpha}$−$^{13}C^{\beta}$ pairs of a number of Ala residues in MSG makes the recording of ^{13}CO chemical shifts indispensible for unambiguous and reliable assignments. Practically complete assignments of Ala methyls in MSG have been obtained using either a 4D ALA-HMCBCACO experiment or a combination of 3D HMCBCA(CO) and 3D HMCB(CA)CO data sets. The alanine region of a 2D methyl ^{1}H−^{13}C correlation map of {U−^{13}C}-Pyruvate-derived MSG "filtered" for methyl groups of the $^{13}CH_3$ variety is shown in Figure 2a. Figure 2b−c show selected 2D $^{1}H^{\beta}$−$^{13}C^{\alpha}$ slices from the 4D ALA-HMCBCACO data set. The ^{1}H and ^{13}C chemical shifts of alanine methyls in MSG are listed in the Supporting Information along with a number of remaining ambiguities in methyl assignments arising primarily from degeneracy of *all* the ^{13}C and ^{13}CO chemical shifts observed in several alanine residues. On one occasion (A79 and A589, both located in the beginning of α-helices in the structure of MSG[22]), the similarity of structural environment "seen" by alanine residues led to apparent overlap in all the four dimensions. Of note, the two most upfield-shifted Ala methyl peaks in the ^{1}H dimension belong to A276 and A585 with $^{1}H^{\beta}(^{13}C^{\beta})$ chemical shifts of −0.68(15.6) and 0.68(17.3) ppm, respectively (outside of the spectral window of Figure 2a), a direct consequence of the ring-current effect induced by W610 and W587 indole rings, respectively.

We anticipate that the developed methodology will promote the use of alanine $^{13}CH_3$ methyls as probes of molecular structure and dynamics in large proteins. The NMR experiment described here is expected to be applicable to protein systems with a high degree of chemical shift degeneracy, such as membrane or (partially) unfolded proteins as well as large protein−protein complexes, irrespective of the labeling strategy chosen to selectively protonate alanine methyl sites.

Acknowledgment. This work was supported in part by the Nano-Biotechnology Award (University of Maryland) to V.T. The authors thank Prof. Lewis Kay (University of Toronto) for many useful suggestions.

Supporting Information Available: One table listing $^{1}H^{\beta}$, $^{13}C^{\beta}$, $^{13}C^{\alpha}$, and ^{13}CO chemical shifts of alanines in MSG (37 °C; pH = 7.1). One table listing acquisition parameters of the 4D ALA-HMCMCACO experiment. Listed also is the Bruker pulse code for the 4D ALA-HMCMCACO experiment (Figure 1). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. *J. Biomol. NMR* **1999**, *13*, 369–374.
(2) Tugarinov, V.; Kay, L. E. *J. Biomol. NMR* **2004**, *28*, 165–172.
(3) Tugarinov, V.; Hwang, P. M.; Ollerenshaw, J. E.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 10420–10428.
(4) Sprangers, R.; Kay, L. E. *Nature* **2007**, *445*, 618–622.
(5) Hamel, D. J.; Dahlquist, F. W. *J. Am. Chem. Soc.* **2005**, *127*, 9676–7.
(6) Sprangers, R.; Gribun, A.; Hwang, P. M.; Houry, W. A.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16678–16683.
(7) Velyvis, A.; Yang, Y. R.; Schachman, H. K.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8815–20.
(8) McCaldon, P.; Argos, P. *Proteins* **1988**, *4*, 99–122.
(9) Fernandez, A.; Scott, L. R.; Scheraga, H. A. *J. Phys. Chem.* **2003**, *107*, 9929–9932.
(10) Tugarinov, V.; Muhandiram, R.; Ayed, A.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 10025–10035.
(11) Isaacson, R. L.; Simpson, P. J.; Liu, M.; Cota, E.; Zhang, X.; Freemont, P.; Matthews, S. *J. Am. Chem. Soc.* **2007**, *129*, 15428–15429.
(12) Gardner, K. H.; Rosen, M. K.; Kay, L. E. *Biochemistry* **1997**, *36*, 1389–1401.
(13) Rosen, M. K.; Gardner, K. H.; Willis, R. C.; Parris, W. E.; Pawson, T.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 627–636.
(14) Venters, R. A.; Farmer, B. T.; Fierke, C. A.; Spicer, L. D. *J. Mol. Biol.* **1996**, *264*, 1101–1116.
(15) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 13868–78.
(16) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335–338.
(17) Shaka, A. J.; Lee, C. J.; Pines, A. *J. Magn. Reson.* **1988**, *64*, 547.
(18) Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1990**, *89*, 496–514.
(19) Boyd, J.; Soffe, N. *J. Magn. Reson.* **1989**, *85*, 406–413.
(20) Geen, H.; Freeman, R. *J. Magn. Reson.* **1991**, *93*, 93–141.
(21) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393–399.
(22) Howard, B. R.; Endrizzi, J. A.; Remington, S. J. *Biochemistry* **2000**, *39*, 3156–3168.

JA808202Q