Production and Incorporation of ^{15}N, ^{13}C, ^{2}H (^{1}H-δ1 Methyl) Isoleucine into Proteins for Multidimensional NMR Studies

Kevin H. Gardner* and Lewis E. Kay

Protein Engineering Network Centres of Excellence and Departments of Medical Genetics and Microbiology Biochemistry and Chemistry, University of Toronto Toronto, Ontario, Canada, M5S 1A8

Received March 3, 1997
Revised Manuscript Received May 27, 1997

The production of macromolecules labeled with NMR-active isotopes has greatly increased the range of systems that are amenable to multidimensional solution NMR studies. Uniform ^{15}N, ^{13}C labeling has facilitated the development of triple resonance (^{15}N, ^{13}C, ^{1}H) experiments, making possible detailed structural and dynamics studies of systems up to approximately 20 kDa molecular mass.[1] NMR spectra of larger macromolecules are often complicated by both increasing numbers of cross-peaks and a concomitant growth in line widths that decreases signal-to-noise and resolution. These problems can be significantly reduced by deuterating aliphatic carbon sites, improving the sensitivity and resolution of many experiments that rely on magnetization transfer through ^{13}C nuclei.[2−7]

Unfortunately, deuteration also removes many of the protons that are the source of nuclear Overhauser effect (NOE)-based distance restraints that are essential to determining high-resolution structures. With this in mind, we have developed a strategy whereby proteins are overexpressed in bacteria grown in $^{2}H_2O$ with (^{13}C, ^{1}H)-pyruvate as the sole carbon source.[8] Such proteins are highly deuterated in virtually all aliphatic positions while retaining between 40−80% of the methyl protons in four amino acids: Ala, Val, Leu, and Ile (γ2 only). Chemical shift assignments of backbone and methyl group nuclei are obtained using recently developed pulse schemes,[3,4,9,10] while methyl−methyl, methyl-NH, and NH-NH distance restraints are established from a series of 4D NOE experiments.[11−14] Incorporating distance restraints from these methyl-based NOEs improves the precision and accuracy of structures relative to those generated using NOEs solely between backbone amide protons.[15−17]

However, the methyl groups of several amino acids are not protonated using this pyruvate-based strategy,[8] including the isoleucine δ1 methyl group. Protonation at the Ile δ1 methyl position is desirable since cross-peaks from these methyl groups are often significantly better resolved than signals from other methyl types in ^{13}C−^{1}H shift correlation spectra, even in proteins larger than 40 kDa. This is illustrated in a ^{13}C−^{1}H correlation spectrum of a 41.0 kDa maltose binding protein (MBP)−maltose complex (Figure 1). Additionally, in contrast to valine γ and leucine δ methyl groups, the lone isoleucine δ1 methyl does not have to be stereospecifically assigned.

From a structural perspective, isoleucine is an important residue because of its enrichment in protein hydrophobic cores.[18] The location of the isoleucine δ1 Me group at a substantial distance (approximately 3.7 Å) from the protein backbone facilitates the observation of NOEs between Ile δ1 methyl protons and amino acid side chains that are distant in the primary sequence. In an analysis of over 290 nonhomologous protein structures solved to better than 2.5 Å resolution,[19] we have found an average of 5.1 ± 1.9 backbone amide protons and a total of 2.8 ± 1.5 Ile δ1, Val γ, and Leu δ Me groups within 6.0 Å of a given Ile δ1 Me group. Methyl−methyl NOEs are particularly important for structure determination as they involve residues with a median separation of 30 amino acids, as opposed to two and three for amide−amide and amide−methyl NOEs, respectively.

Our approach to generate uniformly (^{15}N, ^{13}C)-labeled, highly deuterated, δ1 methyl-protonated isoleucine follows the biosynthetic route of *Escherichia coli* starting with ^{15}N, ^{13}C, ^{1}H threonine (Scheme 1). Isoleucine is produced by growing prototrophic *E. coli* in a $^{2}H_2O$-based minimal medium with $^{15}NH_4Cl$ and 50 mg/L of (3,3-$^{2}H_2$) ^{13}C 2-ketobutyrate as generated in steps 1 and 2 (details in Supporting Information). In this particular case, (^{12}C, ^{1}H) pyruvate was used as the carbon source, as previous studies have established that pyruvate does not efficiently protonate isoleucine δ1 positions.[8] As such, any protonation at these sites is derived from the added 2-ketobutyrate. In practice, to generate uniformly ^{13}C labeled proteins with a highly deuterated background, a carbon source such as (^{13}C, ^{2}H) glucose would be used. Additionally, we demonstrate that further supplementing the media with 50 mg/L of (^{15}N, ^{13}C, ^{1}H) valine results in the production of uniformly ^{13}C, partially deuterated, fully methyl-^{1}H labeled valine and leucine in overexpressed proteins.[20,21]

Figure 2 presents a ^{13}C−^{1}H constant time HSQC spectrum[22,23] recorded on a sample of the 105 amino acid C-terminal SH2 domain of bovine phospholipase Cγ1 (PLCγ1) generated in this manner. Only the CH_3 isotopomers of the Ile δ1, Val γ, and Leu δ methyl groups are observed, while no other methyl positions are protonated. Quantitation of HNCO-based spectra demonstrate that 92 ± 5% of the valine and the leucine methyl groups and essentially 100% of the isoleucine δ1 methyl groups are fully (^{13}C, ^{1}H) labeled. The absence of ^{13}C−^{1}H correlations at most other valine, leucine, and isoleucine positions indicates that these sites are highly deuterated. Two exceptions to this are the Val β and Leu γ methine positions, which both show

* Corresponding author. E-mail address: gardner@bloch.med.utoronto.ca.

(1) Clore, G. M.; Gronenborn, A. M. *Methods Enzymol.* **1994**, *239*, 349−363.
(2) Grzesiek, S.; Anglister, J.; Ren, H.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 4369−4370.
(3) Yamazaki, T.; Lee, W.; Arrowsmith, C. H.; Muhandiram, D. R.; Kay, L. E. *J. Am. Chem. Soc.* **1994**, *116*, 11655−11666.
(4) Yamazaki, T.; Lee, W.; Revington, M.; Mattiello, D. L.; Dahlquist, F. W.; Arrowsmith, C. H.; Kay, L. E. *J. Am. Chem. Soc.* **1994**, *116*, 6464−6465.
(5) Nietlispach, D.; Clowes, R. T.; Broadhurst, R. W.; Ito, Y.; Keeler, J.; Kelly, M.; Ashurst, J.; Oschkinat, H.; Domaille, P. J.; Laue, E. D. *J. Am. Chem. Soc.* **1996**, *118*, 407−415.
(6) Farmer II, B. T.; Venters, R. A. *J. Am. Chem. Soc.* **1995**, *117*, 4187−4188.
(7) Farmer II, B. T.; Venters, R. A. *J. Biomol. NMR* **1996**, *7*, 59−71.
(8) Rosen, M. K.; Gardner, K. H.; Willis, R. C.; Parris, W. E.; Pawson, T.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 627−636.
(9) Shan, X.; Gardner, K. H.; Muhandiram, D. R.; Rao, N. S.; Arrowsmith, C. H.; Kay, L. E. *J. Am. Chem. Soc.* **1996**, *118*, 6570−6579.
(10) Gardner, K. H.; Konrat, R.; Rosen, M. K.; Kay, L. E. *J. Biomol. NMR* **1996**, *8*, 351−356.
(11) Grzesiek, S.; Wingfield, P.; Stahl, S.; Kaufman, J. D.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 9594−9595.
(12) Muhandiram, D. R.; Xu, G. Y.; Kay, L. E. *J. Biomol. NMR* **1993**, *3*, 463−470.
(13) Venters, R. A.; Metzler, W. J.; Spicer, L. D.; Mueller, L.; Farmer II, B. T. *J. Am. Chem. Soc.* **1995**, *117*, 9592−9593.
(14) Vuister, G. W.; Clore, G. M.; Gronenborn, A. M.; Powers, R.; Garrett, D. S.; Tschudin, R.; Bax, A. *J. Magn. Reson. B* **1993**, *101*, 210−213.
(15) Metzler, W. J.; Wittekind, M.; Goldfarb, V.; Mueller, L.; Farmer II, B. T. *J. Am. Chem. Soc.* **1996**, *118*, 6800−6801.
(16) Smith, B. O.; Ito, Y.; Raine, A.; Teichmann, S.; Ben-Tovim, L.; Nietlispach, D.; Broadhurst, R. W.; Terada, T.; Kelly, M.; Oschkinat, H.; Shibata, T.; Yokoyama, S.; Laue, E. D. *J. Biomol. NMR* **1996**, *8*, 360−368.
(17) Gardner, K. H.; Rosen, M. K.; Kay, L. E. *Biochemistry* **1997**, *36*, 1389−1401.
(18) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155−164.
(19) Heringa, J.; Sommerfeldt, H.; Higgins, D.; Argos, P. *CABIOS* **1992**, *8*, 599−600.
(20) Crespi, H. L.; Katz, J. J. *Nature* **1969**, *224*, 560−562.
(21) Oda, Y.; Nakamura, H.; Yamazaki, T.; Nagayama, K.; Yoshida, M.; Kanaya, S.; Ikehara, M. *J. Biomol. NMR* **1992**, *2*, 137−147.
(22) Vuister, G. W.; Bax, A. *J. Magn. Reson.* **1992**, *98*, 428−435.
(23) Santoro, J.; King, G. C. *J. Magn. Reson.* **1992**, *97*, 202−207.

S0002-7863(97)00651-3 CCC: $14.00 © 1997 American Chemical Society

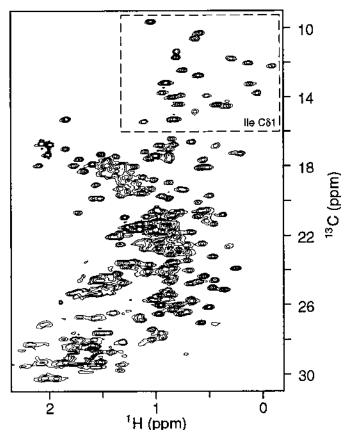
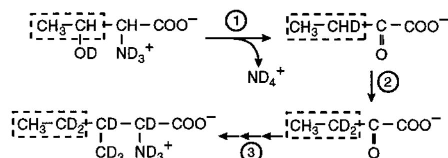


Figure 1. $^{13}C-^{1}H$ HSQC spectrum of a 41 kDa maltose binding protein/maltose complex. The boxed area designates the region containing cross-peaks from Ile $\delta 1$ methyl groups. The sample was 2.2 mM in maltose binding protein (natural abundance ^{13}C) with 1.8 equiv of added maltose in 20 mM sodium phosphate buffer (pH 7.2), 3 mM sodium azide, and 10% $^{2}H_2O$. 192 scans/FID were acquired in 20 h total acquisition time at 37 °C on a Varian Inova 500 MHz spectrometer.

Scheme 1. Method Used To Generate ^{13}C, ^{2}H (^{1}H-Methyl) Isoleucine[a]



[a] Step 1: Conversion from threonine (20 mM) into (3-^{2}H) 2-ketobutyrate, catalyzed by *E. coli* biosynthetic threonine deaminase[30] in 99.5% $^{2}H_2O$, T = 26 °C, 4 h. Step 2: Conversion of (3-^{2}H) 2-ketobutyrate (diluted to ~2.7 mM) into (3,3-$^{2}H_2$) 2-ketobutyrate by proton/deuterium exchange at C3 using pH* (uncorrected) 10.2, 99.5% $^{2}H_2O$, $T = 45$ °C, 13 h. Step 3: Conversion of (3,3-$^{2}H_2$) 2-ketobutyrate into isoleucine and eventual incorporation into overexpressed protein, carried out by *E. coli* metabolism (protein expressed in 99.5% $^{2}H_2O$, T = 37 °C).

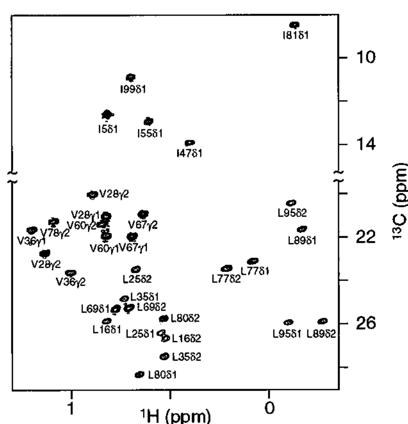


Figure 2. $^{13}C-^{1}H$ constant time HSQC spectrum of a highly deuterated, methyl-protonated C-terminal SH2 domain of bovine PLCγ1, labeled using the approach of Scheme 1 and including 50 mg/L of ^{15}N, ^{13}C valine in the growth media. The sample was approximately 200 μM protein in 100 mM sodium phosphate buffer (pH 6.0), 100 μM EDTA, 100 μM DTT, and 10% $^{2}H_2O$. 64 scans/FID were acquired (8 hr total acquisition time) at 30 °C on a Varian Inova 600 MHz spectrometer.

weak cross-peaks. Residual protonation at these sites can be reduced by using minimal media supplemented with (2,3-$^{2}H_2$) ^{15}N, ^{13}C-valine (synthesized[24] or purchased from commercial sources). To summarize, when valine labeled in this manner is used together with (3,3-$^{2}H_2$) ^{13}C 2-ketobutyrate, (^{13}C, ^{2}H) glucose, and $^{15}NH_4Cl$ in a D_2O-based medium, it is possible to generate a ^{15}N, ^{13}C labeled protein with fully protonated Val γ, Leu δ, and Ile $\delta 1$ methyl groups with high levels of deuteration at other aliphatic positions.

The impact of distance restraints from the methyl groups of Val, Leu, and Ile ($\delta 1$) on the precision and accuracy of NMR-derived structures was assessed by calculating structures of the C-terminal SH2 domain from PLCγ1 using several sets of simulated structural restraints. A list of all possible NOEs was generated by choosing pairs of NH-NH, NH-methyl carbon, and methyl carbon-methyl carbon groups in a distance- and type-dependent manner,[17] and five distance-weighted random subsets of this group were selected, each containing approximately 240 restraints. These were combined with 72 conservative backbone dihedral angle (ϕ,ψ) restraints and input into X-PLOR[25] to obtain a set of 28 structures that had good agreement with experimental and empirical restraints. This group exhibited a precision (pairwise-to-mean) of 2.23 ± 0.41 Å and accuracy (pairwise to a high resolution, ref 26) of 3.20 ± 0.72 Å through all residues found in secondary structure elements as judged by ^{13}C chemical shifts.[27] These results compare quite favorably to those previously found using a larger number of distance restraints from all of the pyruvate-protonated methyl groups and are significantly better than statistics obtained for structures generated using distance restraints only from backbone NH atoms.[17]

In summary, we have described a scheme for the production of proteins that are uniformly ^{15}N, ^{13}C labeled and extensively deuterated except for the methyl groups of Val, Leu, and Ile ($\delta 1$ only). This method offers several advantages over previously suggested alternatives for producing highly deuterated, site-protonated proteins[15,16,20,21,28,29] in that these methods typically result in fully protonated side chains with significant deuteration only at the Cα position. The high extent of deuteration with the present protocol ensures that the experiments developed for backbone and side-chain assignment of deuterated proteins will be of high sensitivity. The uniformly protonated methyl groups on the three amino acids that constitute over 35% of protein hydrophobic cores[18] allows for the assignment of long-range NOE-based interproton distance restraints for structure determination. This labeling scheme will play an important role in facilitating NMR-based structural studies of proteins in the 30-50 kDa range.

Acknowledgment. The authors thank Dr. Edward Eisenstein, Dr. Diana Chinchilla, and Kathryn Fisher (CARB, University of Maryland) for *E. coli* threonine deaminase and advice regarding expression of this enzyme, Dr. Kalle Gehring (BRI, Montreal) for providing an MBP expression plasmid and advice on its purification, Mr. Randall Willis (Hospital for Sick Children, Toronto) for assistance with protein purification, Dr. Michael Rosen (Sloan Kettering, New York) for valuable discussions, and Dr. Yves Aubin (Merck Frosst, Montreal) for a critical evaluation of the manuscript. This work was supported by grants from the Medical Research Council of Canada and a postdoctoral fellowship to K.H.G. from the Helen Hay Whitney Foundation.

Supporting Information Available: A listing of reaction conditions for all steps of Scheme 1, ^{1}H NMR spectra monitoring the progress of steps 1 and 2, spectral parameters for Figures 1 and 2, and quantitation of methyl labeling efficiency (5 pages). See any current masthead page for ordering and Internet access instructions.

JA9706514

(24) Chanatry, J. A.; Schafer, P. H.; Kim, M. S.; LeMaster, D. M. *Anal. Biochem.* **1993**, *213*, 147-151.
(25) Brünger, A. T. *X-PLOR 3.1 Manual*; Yale University Press: New Haven, 1992.
(26) Pascal, S. M.; Singer, A. U.; Kay, L. E.; Forman-Kay, J. D. manuscript in preparation.
(27) Wishart, D. S.; Sykes, B. D. *J. Biomol. NMR* **1994**, *4*, 171-180.
(28) Brodin, P.; Drakenberg, T.; Thulin, E.; Forsén, S.; Grundström, T. *Prot. Eng.* **1989**, *2*, 353-358.
(29) Markley, J. L.; Putter, I.; Jardetzky, O. *Science* **1968**, *161*, 1249-1251.
(30) Eisenstein, E. *J. Biol. Chem.* **1991**, *266*, 5801-5807.