An NMR Experiment for Measuring Methyl−Methyl NOEs in ^{13}C-Labeled Proteins with High Resolution

Catherine Zwahlen,[†,‡] Kevin H. Gardner,[†] Siddhartha P. Sarma,[§] David A. Horita,[§] R. Andrew Byrd,*[,§] and Lewis E. Kay*[,†]

Contribution from the Protein Engineering Network Centers of Excellence and Departments of Molecular and Medical Genetics, Biochemistry and Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A8, Hospital for Sick Children, Biochemistry Research, Toronto, Ontario, Canada M5G 1X8, and Macromolecular NMR Section, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702-1201

Received April 9, 1998

Abstract: A three-dimensional NMR experiment is presented for correlating NOEs between methyl groups in protonated, ^{15}N,^{13}C-labeled or methyl protonated, highly deuterated ^{15}N,^{13}C-labeled proteins. The high resolution of this experiment facilitates the assignment of NOEs between methyls that can be poorly resolved in other three- and four-dimensional experiments and extends the utility of solution-based NMR structural studies to proteins in the 40 kDa molecular weight regime. Applications to methyl protonated, highly deuterated samples of the 370 residue maltose binding protein (122 methyl groups) and the dimer of the 124 residue amino terminal domain of human STAT-4 (59 methyls) are presented. The method is also well suited for studies of fully protonated proteins, as demonstrated with an application involving the 160 residue phosphotyrosine binding domain from *Drosophila* Numb. Distance restraints obtained from the present experiment are particularly useful in the generation of global protein folds since many methyls are located in hydrophobic protein cores and inter-methyl restraints therefore link elements of secondary structure that are often distant in the primary sequence.

Introduction

The development of multidimensional, multinuclear NMR spectroscopy over the past decade has significantly enlarged the scope of problems that are now studied by NMR approaches.[1,2] Assignment of resonances in ^{15}N,^{13}C-labeled or ^{15}N,^{13}C,^{2}H-labeled proteins of sizes on the order of 20−30 kDa or even larger can often be accomplished in relatively short order with use of a large number of powerful experiments. With the advances in techniques for resonance assignment, the most difficult and time-consuming portion of the structure determination processes is the assignment of NOE cross-peaks to specific, proximal proton pairs. In this regard the development of ^{13}C-edited three-[3] and four-dimensional[4−6] NOE experiments has been particularly useful. Nevertheless, many proteins of biological importance are of a size and complexity such that solution structure determination by existing NMR methods remains difficult. For example, overlap in NOE spectra of even modestly sized proteins is significant and complicates assignment of these multidimensional spectra. This is a major limitation, since in the absence of a critical number of distance correlations resultant structures are both imprecise and inaccurate.[7]

In principle, there are at least two approaches for overcoming the problem of ambiguous distance restraints arising from overlap of cross-peaks in NOE spectra. Recently Nilges et al. have described an elegant computational method that makes use of an iterative automated assignment strategy for the interpretation of ambiguous restraints.[8,9] In our laboratories we have applied this methodology with success to the solution structure determination of a peptide complex of a 160-residue phosphotyrosine binding domain (PTB) from *Drosophila* Numb (dNumb) and to the NusB protein (139 residues). An alternative and complementary approach is to design additional NMR experiments, often in combination with novel isotopic labeling strategies,[10−12] which provide improved resolution in the hopes of assigning a larger fraction of NOEs prior to structure calculation. To this end we have recently described an experiment for measuring NOE correlations involving protons

* Address correspondence to these authors.
† University of Toronto.
‡ Hospital for Sick Children.
§ NCI-Frederick Cancer Research and Development Center.

(1) Bax, A. *Curr. Opin. Struct. Biol.* **1994**, *4*, 738−744.
(2) Kay, L. E.; Gardner, K. H. *Curr. Opin. Struct. Biol.* **1997**, *7*, 722−731.
(3) Ikura, M.; Kay, L. E.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1990**, *86*, 204−209.
(4) Vuister, G. W.; Clore, G. M.; Gronenborn, A. M.; Powers, R.; Garrett, D. S.; Tschudin, R.; Bax, A. *J. Magn. Reson. Ser. B* **1993**, *101*, 210−213.
(5) Clore, G. M.; Kay, L. E.; Bax, A.; Gronenborn, A. M. *Biochemistry* **1991**, *30*, 12−18.
(6) Zuiderweg, E. R. P.; Petros, A. M.; Fesik, S. W.; Olejniczak, E. T. *J. Am. Chem. Soc.* **1991**, *113*, 370−371.
(7) Gardner, K. H.; Rosen, M. K.; Kay, L. E. *Biochemistry* **1997**, *36*, 1389−1401.
(8) Nilges, M. *J. Mol. Biol.* **1995**, *245*, 645−660.
(9) Nilges, M.; Macias, M. J.; O'Donoghue, S. I.; Oschkinat, H. *J. Mol. Biol.* **1997**, *269*, 408−422.
(10) Gardner, K. H.; Kay, L. E. *J. Am. Chem. Soc.* **1997**, *119*, 7599−7600.
(11) Smith, B. O.; Ito, Y.; Raine, A.; Teichmann, S.; Ben-Tovim, L.; Nietlispach, D.; Broadhurst, R. W.; Terada, T.; Kelly, M.; Oschkinat, K.; Shibata, T.; Yokoyama, S.; Laue, E. D. *J. Biomol. NMR* **1996**, *8*, 360−368.
(12) Metzler, W. J.; Wittekind, M.; Goldfarb, V.; Mueller, L.; Farmer, B. T. *J. Am. Chem. Soc.* **1996**, *118*, 6800−6801.

S0002-7863(98)01205-0 CCC: $15.00
Published on Web 07/18/1998

from C^{γ} methyl groups of Val and Ile in ^{15}N,^{13}C-labeled or methyl-protonated, ^{15}N,^{13}C,^{2}H-labeled proteins.[13] Schemes for measuring NOEs connecting methyl groups are particularly important since methyls are often in the hydrophobic cores of proteins and bring together elements of secondary structure that are well separated in primary sequence.[14]

In this paper we present an experiment that addresses the sensitivity and resolution requirements for obtaining NOE connectivities between pairs of methyl groups in proteins. The utility of the method is demonstrated by applications to methyl-protonated, ^{15}N,^{13}C,^{2}H-labeled samples of both a 42 kDa complex consisting of maltose binding protein (MBP, 370 residues) and β-cyclodextrin as well as a 29 kDa symmetric dimer comprising the amino-terminal domain (residues 1−124) of the intact human STAT-4 protein[15] (hSTAT4$_{1-124}$, Signal Transducers and Activators of Transcription). In addition to the usefulness of the experiment for applications to systems involving methyl-protonated, highly deuterated proteins, the methodology is also valuable for fully protonated proteins or protein complexes, as illustrated with an application to an ^{15}N,^{13}C-labeled sample of a complex of the dNumb PTB domain[16] (17.9 kDa) with a high affinity binding, unlabeled peptide.

Materials and Methods

A sample of the protonated, ^{15}N,^{13}C-labeled dNumb PTB domain in complex with a high affinity binding seven residue peptide (Ac-AYIGPpYL) was prepared as described previously.[16] Sample conditions were 1.5 mM protein with equimolar peptide, 20 mM sodium phosphate, 150 mM sodium chloride, 0.1 mM benzamidine, 0.1 mM EDTA, 20 mM deuterated DTT, >99% D_2O, and pD 6.0.

^{15}N, ^{13}C, ^{2}H, methyl protonated (Val, Leu and Ile $C^{\delta 1}$) *E. coli* MBP was generated as described previously[10] and a final sample consisting of 1.4 mM protein, 2.0 mM β-cyclodextrin, 20 mM sodium phosphate (pH 7.2), 3 mM NaN_3, 100 mM EDTA, 0.1 mg/mL Pefabloc, 1 μg/μL pepstatin, and 10% D_2O was prepared. As described in some detail elsewhere,[13] omission of a step where the ^{1}H-labeled bacteria are diluted in deuterium rich media coupled with the use of an inappropriately long induction time in the overexpression of this sample resulted in the production of protein with small amounts of CH_2D and CHD_2 methyl isotopomers. Cross-peaks arising from these groups were eliminated with use of the purging scheme described below.

^{15}N, ^{13}C, ^{2}H, methyl protonated (Val, Leu, and Ile $C^{\delta 1}$) hSTAT4$_{1-124}$ was prepared as described by Gardner and Kay.[10] The details of sample production will be given elsewhere. The ^{13}C CT-HSQC[17,18] of hSTAT4$_{1-124}$ showed the expected 59 methyl groups (Val, Leu, and Ile $C^{\delta 1}$) with no evidence of any CH_2D or CHD_2 isotopomers. Sample conditions were 0.7 mM protein (monomer concentration), 20 mM sodium acetate-d_3 (pH 5.3), 50 mM sodium chloride, 5 mM DTT, and 90% H_2O/10% D_2O.

A 3D ^{13}C,^{13}C methyl NOESY spectrum of the fully protonated PTB domain was recorded at 30 °C with use of the scheme of Figure 1b on a Varian 600 MHz Inova spectrometer. A matrix comprised of (78, 30, 448) complex data points in each of (t_1, t_2, t_3) was obtained with spectral widths of (3200 Hz, 3200 Hz, 9000 Hz) in (F_1, F_2, F_3), corresponding to acquisition times of (24.4 ms, 9.4 ms, 50 ms). A mixing time of 90 ms, 48 scans/FID, and a repetition delay of 0.9 s were used, giving rise to a net acquisition time of 123 h.

The pulse sequence of Figure 1a was employed in recording the data set on the MBP/β-cyclodextrin complex at 37 °C (Varian 600 MHz Inova). A matrix consisting of (60, 60, 448) complex points, spectral widths of (2445 Hz, 2445 Hz, 9000 Hz) and acquisition times of (24.5 ms, 24.5 ms, 50 ms) was recorded. A mixing time of 100 ms was used along with 32 scans/FID and a repetition delay of 0.85 s to give a net measuring time of 122 h. A 4D $^{13}C-^{13}C$ NOESY[4] data set was recorded with spectral widths of 3600 Hz in each of F_1, F_2, and F_3 and 9000 Hz in F_4. Acquisition times of (13.9 ms, 4.4 ms, 4.4 ms, 50 ms) in (t_1, t_2, t_3, t_4), a mixing time of 100 ms, 4 scans/FID, and a recycle delay of 1.0 s were used, giving rise to a total collection time of 125 h.

The scheme of Figure 1b was used to record the 3D ^{13}C,^{13}C methyl NOESY spectrum of hSTAT4$_{1-124}$ with a number of modifications. All $^{13}C'$, ^{15}N, and ^{2}H pulses were eliminated and the I-BURP2 selective pulses during the INEPT transfers as well as the RE-BURP pulse during the constant time t_1 period were replaced with high-power rectangular pulses. A data matrix consisting of (50, 32, 512) complex points, spectral widths of (2111.8 Hz, 2111.8 Hz, 9000 Hz), and acquisition times of (23.7 ms, 15.2 ms, 57 ms) was utilized for the experiment performed at 25 °C on a Varian Unity+ 600 MHz spectrometer equipped with a Z-spec triple resonance probe (Nalorac Corp., Martinez, CA). Note that the net acquisition time of 15 ms in t_2 used in this experiment is somewhat longer than optimal because of the evolution of magnetization from the one-bond $^{13}C-^{13}C$ coupling that occurs during this period. In principle, $t_{2,\max} \approx 9.5$ ms followed by linear prediction results in spectra of higher signal-to-noise for the same acquisition time (see below). A mixing time of 175 ms was used along with 64 scans/FID and a 1.0 s recycle delay, resulting in an acquisition time of 147 h.

All data sets were processed with NMRPipe/NMRDraw[19] software and analyzed with either ANSIG[20] or NMRView.[21] In the case of spectra recorded on the PTB domain mirror image linear prediction[22] was employed to extend the t_1 (constant time) domain by a factor of 1.5 and forward−backward linear prediction[23] was used to extend the t_2 domain (nonconstant time) by 50% with use of the procedure described in Kay et al.[24] and in more detail recently.[13] In the case of the MBP data set, the t_1 and t_2 time domains were extended by 50% with use of mirror image linear prediction, while in the case of the spectrum recorded on the hSTAT4$_{1-124}$ protein the t_1 time domain was doubled with mirror image linear prediction. The (t_1, t_2, t_3) time domains were apodized with 90° shifted sine-bell squared, 90° shifted sine-bell squared, and 70° shifted sine-bell squared window functions, respectively, prior to Fourier transformation. The absorptive parts of the resultant data sets consisted of 128 × 128 × 448, 128 × 128 × 448, and 256 × 64 × 512 real points for spectra of the PTB, MBP and hSTAT4$_{1-124}$ proteins, respectively. The 4D data set of the MBP/β-cyclodextrin complex was processed with forward−backward linear prediction in t_2 to double the time domain. Gaussian apodization functions were used in the carbon dimensions to minimize truncation artifacts. The absorptive part of the processed data set consisted of 128 × 64 × 32 × 448 points. Note that for both the 3D and 4D spectra the final data matrices were reduced in size (from the numbers given above) by considering only those regions where correlations were present.

Results and Discussion

Figure 1 illustrates the 3D ^{13}C,^{13}C methyl NOESY pulse sequences that were developed for recording NOEs between methyl groups with high resolution. The pulse schemes consist of elements that have been well described in the literature and only a brief discussion of the magnetization flow is therefore

(13) Zwahlen, C.; Vincent, S. J. F.; Gardner, K. H.; Kay, L. E. *J. Am. Chem. Soc.* **1998**, *120*, 4825−4831.
(14) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155−164.
(15) Baden, H. A.; Sarma, S. P.; Kapust, R. B.; Byrd, R. A.; Waugh, D. S. *J. Biol. Chem.* Submitted for publication.
(16) Li, S.-C.; Songyang, Z.; Vincent, S. J. F.; Zwahlen, C.; Wiley, S.; Cantley, L.; Kay, L. E.; Forman-Kay, J. D.; Pawson, T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7204−7209.
(17) Vuister, G. W.; Bax, A. *J. Magn. Reson.* **1992**, *98*, 428−435.
(18) Santoro, J.; King, G. C. *J. Magn. Reson.* **1992**, *97*, 202−207.
(19) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(20) Kraulis, P. J. *J. Magn. Reson.* **1989**, *84*, 627−633.
(21) Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603−614.
(22) Zhu, G.; Bax, A. *J. Magn. Reson.* **1990**, *90*, 405−410.
(23) Zhu, G.; Bax, A. *J. Magn. Reson.* **1992**, *98*, 192−199.
(24) Kay, L. E.; Ikura, M.; Zhu, G.; Bax, A. *J. Magn. Reson.* **1991**, *91*, 422−428.

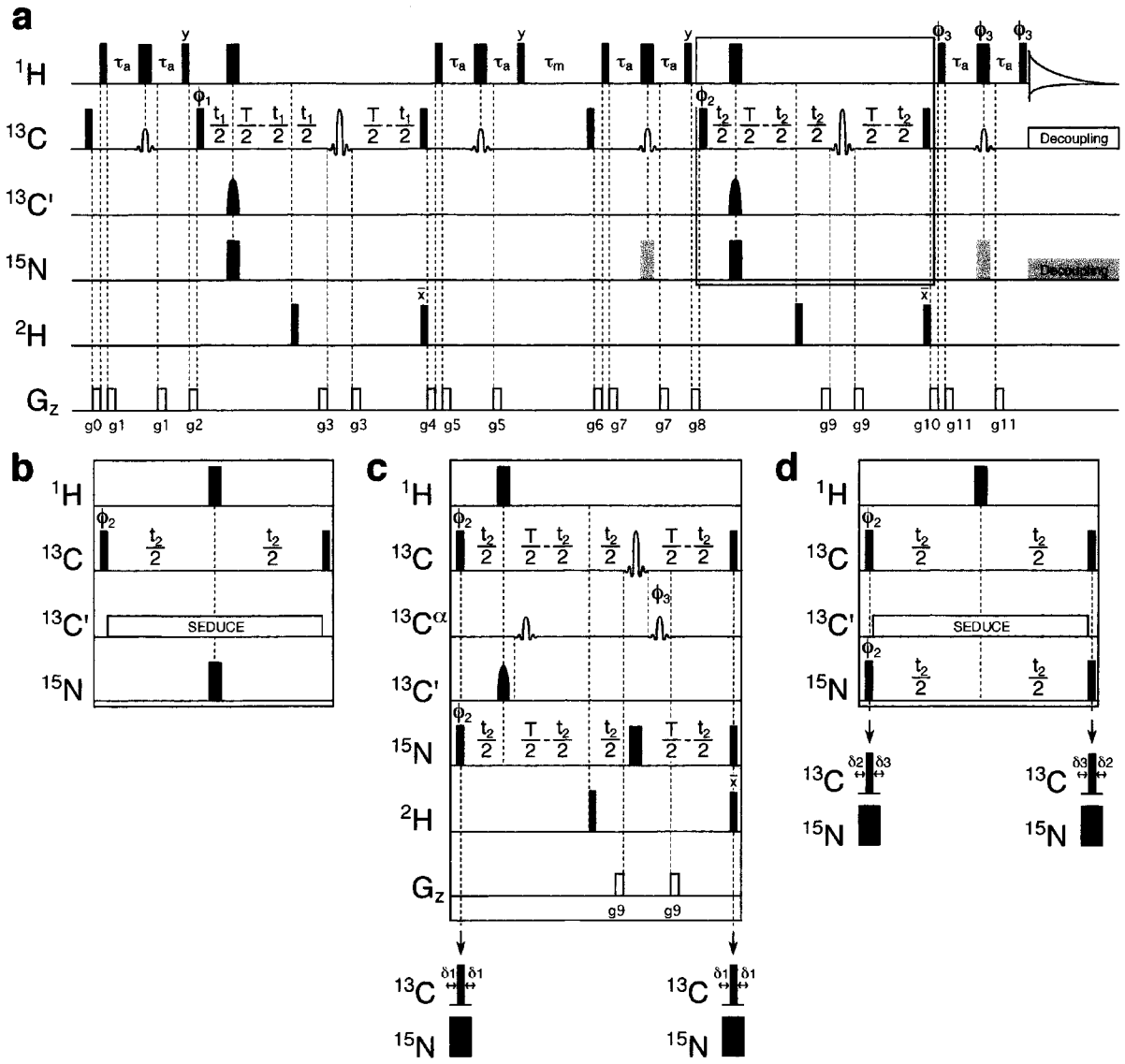

Figure 1. Pulse sequences for recording the 3D ^{13}C,^{13}C methyl NOESY experiment. (a) Double constant time version of the pulse scheme. All narrow (wide) pulses have flip angles of 90° (180°). The ^{1}H, ^{13}C, ^{15}N, and ^{2}H carriers are centered at 4.7 (water), 19.6, 119, and 2.0 ppm, respectively. Rectangular ^{1}H, ^{13}C, ^{15}N, and ^{2}H pulses are applied with field strengths of 30, 21, 5.3, and 1.9 kHz, respectively. All ^{13}C shaped pulses applied during the INEPT[27] periods (of duration $2\tau_a$) are 1.2 ms in length with the I-BURP2[28] profile ($\{1.2 \times 600/y\}$ ms, for experiments recorded at y MHz), while the shaped refocusing pulses applied during the constant time ^{13}C evolution periods (t_1 and t_2 in parts a and c; t_1 in parts b and d) are of 350 μs duration ($\{350 \times 600/y\}$ μs, for experiments recorded at y MHz) and have the RE-BURP[28] profile, with the center of excitation at 25 ppm (5.4 ppm phase modulation of the carrier;[33,34] for applications on fully protonated samples the center of excitation is moved to 35 ppm so that the $^{13}C^{\beta}$ carbons of Thr are properly inverted). The chemical shifts of $^{13}C^{\alpha}$ carbons are within the excitation bandwidth of the RE-BURP pulses. The phases of the RE-BURP pulses are adjusted independently so that for $t_1 = 0$ or $t_2 = 0$ only one of the quadrature components has nonzero signal. The ^{13}C′ pulses have the SEDUCE-1[35] profile and are of 230 μs duration. These pulses are centered at 176 ppm by phase modulation of the carrier.[33,34] For fully protonated samples or methyl protonated, deuterated samples where only the CH_3 isotopomer is present the ^{2}H pulses are eliminated. ^{13}C decoupling during acquisition is achieved with use of a 2.7 kHz WALTZ-16 field.[36] Note that the ^{15}N rectangular pulses indicated by shading are applied only when both ^{15}N and ^{13}C chemical shifts are recorded during t_2 (parts c and d). (b) Single constant time version of the 3D ^{13}C,^{13}C methyl NOESY. ^{13}C′ decoupling is achieved during t_2 with use of a 156.4 ppm cosine-modulated WALTZ-16 scheme,[37] where each element of the profile consists of pulses with the SEDUCE-1 shape. (c) Simultaneous recording of ^{13}C and ^{15}N chemical shifts during t_2 in a constant time mode. The delay $\delta 1$ is set to 0.5(pwn − pwc), where pwn and pwc are the ^{15}N and ^{13}C 90° pulse widths. The $^{13}C^{\alpha}$ selective pulses have the I-BURP2 profile (1.2 ms for applications at 600 MHz; $\{1.2 \times 600/y\}$ ms, for experiments recorded at y MHz) and are centered at 54.5 ppm by phase modulation of the carrier. The second $^{13}C^{\alpha}$ pulse is time inverted relative to the first $^{13}C^{\alpha}$ pulse. (d) Simultaneous recording of ^{13}C and ^{15}N chemical shifts during t_2 in a nonconstant time mode. The delays $\delta 2$ and $\delta 3$ are set to $[(\pi - 2)/\pi]$(pwn − pwc) and $(2/\pi)$(pwn − pwc), respectively. Note that it is possible to adjust the duration of the first t_2 point, $t_2(0)$, so that a 180° first-order phase correction in F_2 for both ^{15}N and^{13}C is obtained.[29] In parts c and d, ^{15}N decoupling during acquisition is achieved with use of a 0.9−1.0 kHz WALTZ-16 field. The delays used are $2\tau_a = 4.5$ ms and $T = 26$ ms, where $2\tau_a$ is the time between successive ^{1}H 90° pulses of the INEPT and includes the time for application of the I-BURP2 pulse. Note that because evolution of magnetization due to the one-bond ^{1}H-^{13}C coupling proceeds at a rate that is smaller than J_{CH} during the simultaneous application of the I-BURP2/^{1}H 180° refocusing pulses, the optimal value of $2\tau_a$ will be greater than $1/(2J_{HC})$. The value of $2\tau_a$ used was obtained by optimizing the signal in 1D spectra. In the case where the I-BURP2 pulses are replaced by high-power rectangular pulses the values of $2\tau_a$ must be adjusted (shortened) accordingly. When both ^{15}N and ^{13}C chemical shifts are recorded during t_2, the INEPT periods flanking the t_2 period can be adjusted to maximize transfer between ^{1}H and ^{13}C and ^{1}HN and ^{15}N, as described previously.[29] However, we prefer to optimize the experiment to achieve maximum sensitivity for methyl−methyl correlations. Quadrature in F_1 and F_2 is obtained by STATES-TPPI[38] of $\phi 1$ and $\phi 2$, respectively. The phase cycling employed is $\phi 1 = 2(x),2(-x)$; $\phi 2 = 4(x),4(-x)$; $\phi 3 = x,-x$; Acq $= x,2(-x),x,-x,2(x),-x$. The duration and strengths of the gradients are as follows: g0 = (500 μs, 10 G/cm); g1 = (400 μs, 5 G/cm); g2 = (1 ms, 15 G/cm); g3 = (400 μs, 18 G/cm); g4 = (800 μs, 15 G/cm); g5 = (500 μs, 6 G/cm); g6 = (1 ms, 8 G/cm); g7 = (200 μs, 7 G/cm); g8 = (500 μs, 10 G/cm); g9 = (400 μs, 25 G/cm); g10 = (800 μs, −15 G/cm); g11 = (700 μs, 8 G/cm). In the case where ^{2}H pulses are applied the ^{2}H lock channel is interrupted at the start of the sequence and re-engaged at the end.

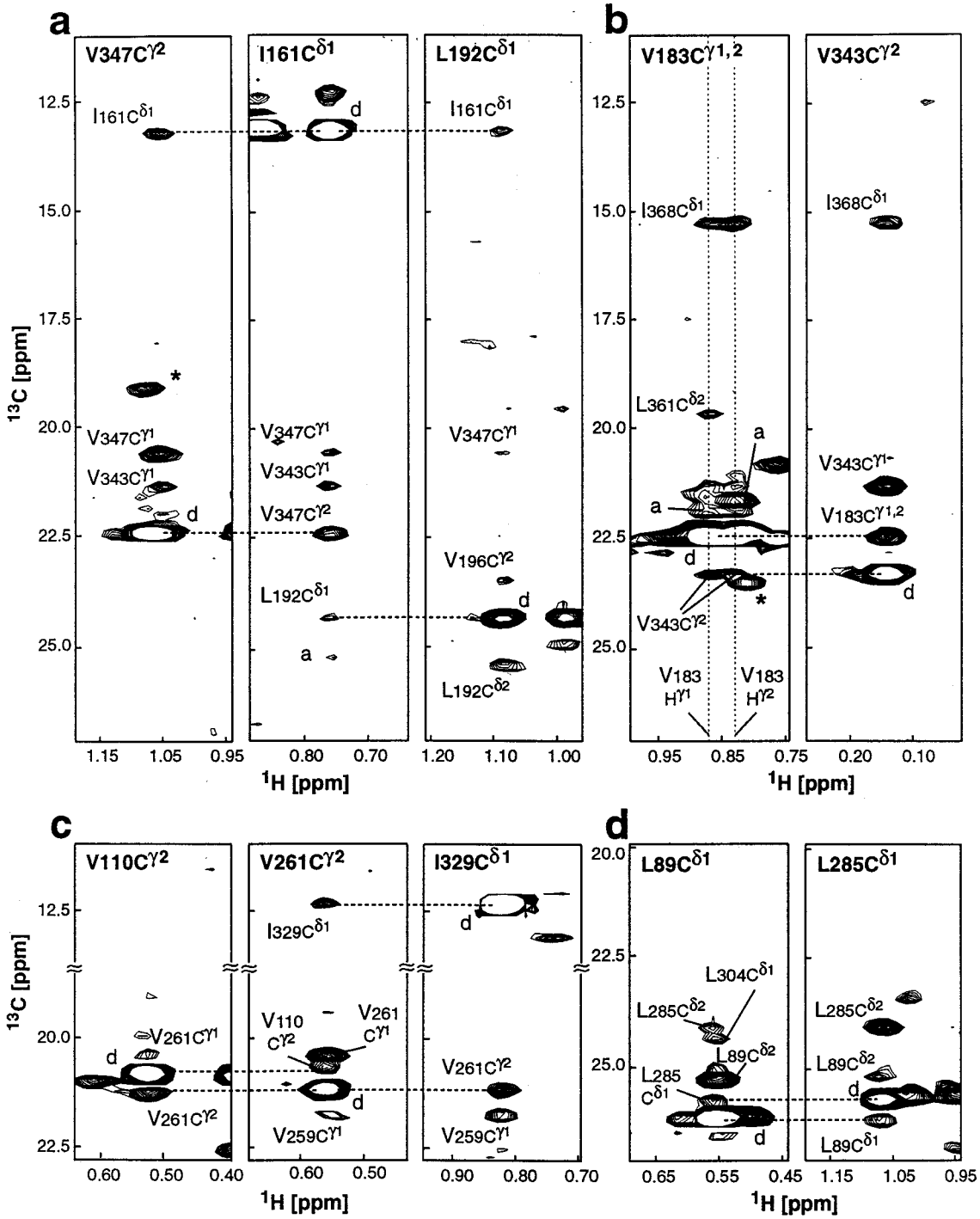

Figure 2. F_1-F_3 strips from the 3D $^{13}C,^{13}C$ methyl NOESY spectrum recorded on the methyl-protonated-(Val, Leu, Ile C^{δ}), $^{15}N,^{13}C$, ^{2}H sample of the MBP/β-cyclodextrin complex (mixing time 100 ms) with the pulse scheme of Figure 1a. NOE correlations linking residues *i* and *j* are observed at ($\omega_{Ci},\omega_{Cj},\omega_{Hj}$); the identity of the destination methyl carbon (C*j*) is indicated in the top of each trace. Diagonal peaks are marked "d", ambiguous peaks are indicated by "a", and cross-peaks with greater intensity in adjacent slices are marked with an asterisk. Symmetry related correlations are connected by dashed horizontal lines.

provided. The transfer steps can be summarized by,

$$^{1}H_i \xrightarrow{J_{HC}} {}^{13}C_i\,(CT\ t_1) \xrightarrow{J_{HC}} {}^{1}H_i \xrightarrow{NOE} {}^{1}H_j \xrightarrow{J_{HC}} {}^{13}C_j\,(CT\ t_2 \text{ or non-CT } t_2) \xrightarrow{J_{HC}} {}^{1}H_j\,(t_3) \quad (1)$$

where the active couplings involved in each transfer are indicated above the arrows and t_n ($n = 1, 2, 3$) is an evolution period. Note that the t_1 period is recorded in constant time mode[17,18] (CT t_1), with the constant time delay, *T*, set to the inverse of the one-bond aliphatic−aliphatic coupling constant, $1/J_{CC}$ ($\approx$26 ms). Thus, evolution due to J_{CC} during the *T* delay does not attenuate the signal. The second carbon evolution period can be recorded in either a constant time or nonconstant time mode, depending on resolution and sensitivity considerations (see below). Fourier transformation of the resultant data set yields cross-peaks at ($\omega_{Ci},\omega_{Cj},\omega_{Hj}$), $i \neq j$, and diagonal peaks at ($\omega_{Ci},\omega_{Ci},\omega_{Hi}$). Note that a related experiment for recording NOEs between amide protons in ^{15}N-labeled proteins has been described in the literature.[25,26]

To limit the magnetization flow to transfers involving methyl groups in cases where protonation in the sample is not restricted to methyls, the ^{13}C 180° pulses employed during the INEPT[27] transfer periods are applied with the I-BURP2[28] shape profile (1.2 ms). Thus, with the carbon carrier frequency at 19.6 ppm, the inversion profile (≥97%) of the I-BURP2 pulse extends from 8.6 to 30.6 ppm, while resonances downfield of 37.5 ppm are not affected (600 MHz ^{1}H frequency). Note that because there is some overlap of chemical shifts of methyl carbons with other carbons, NOEs involving a limited number of non-methyl protons, including Arg H$^\gamma$ and Lys H$^\delta$, can also be observed in applications on protonated proteins. Finally, for applications to methyl-protonated, highly deuterated proteins where protonation of the methyl positions is not complete, the ^{2}H purge pulses applied during the constant time ^{13}C evolution periods ensure that magnetization only from $^{13}CH_3$ isotopomers is retained. In this regard it is worth mentioning that only $^{13}CH_3$ methyl groups could be detected in high signal-to-noise spectra of hSTAT4$_{1-124}$, while in the case of the MBP sample small levels of $^{13}CH_2D$ and $^{13}CHD_2$ isotopomers were observed. The origin of these isotopomers was traced to a number of errors during protein production (see above), and it is noteworthy that only the $^{13}CH_3$ isotopomer was observed in spectra recorded on previous samples of the protein.

As described above, the second evolution period (t_2) can be recorded in a constant time or a nonconstant time mode. Clearly, resolution is significantly improved by maximizing the acquisition time, albeit at the expense of sensitivity. For equal net measuring times, the gain in sensitivity for the experiment in which magnetization is not recorded in a constant time mode during t_2 relative to the double constant time scheme is given by G, where

$$G = (1/\mathrm{AT})\int_{\mathrm{AT}} \cos(\pi J_{\mathrm{CC}} t_2)\exp(-t_2/T_2)\,\mathrm{d}t_2/\exp(-T/T_2) \quad (2.1)$$

$$= (T_2/\mathrm{AT})[1/(1+(\omega T_2)^2)]\{\omega T_2 \exp(-\mathrm{AT}/T_2)\sin(\omega \mathrm{AT}) + [1-\exp(-\mathrm{AT}/T_2)\cos(\omega \mathrm{AT})]\}/\exp(-T/T_2) \quad (2.2)$$

In eq 2, AT is the total acquisition time in the t_2 dimension of the nonconstant time experiment, J_{CC} is the one-bond aliphatic ^{13}C−^{13}C homonuclear scalar coupling constant, T_2 is a time constant that describes the decay of ^{13}C transverse magnetization during the evolution period, $\omega = \pi J_{CC}$, and T is the duration of the constant time period. The value of T_2 includes contributions from the spin flips of the methyl protons and from the transverse relaxation of in-phase methyl carbon magnetization. Assuming values of (9.5 ms, 35 Hz, 40 ms, 26 ms) for (AT, J_{CC}, T_2, T), a value of G =1.4 is obtained. Note that the expression for G in eq 2 does assume that AT $\lesssim {}^1/_2 J_{CC}$, so that doublet components derived from the ^{13}C−^{13}C coupling are not resolved. All other things being equal, the value of G will of course decrease with increasing T_2. Although larger values of G could be obtained for shorter values of AT, the sensitivity gains would come at the expense of the already limited resolution in the nonconstant time version of the experiment. The above calculation does not include the effects of strong coupling between C$^\gamma$ and C$^\delta$ carbons which may be present in the case of Leu residues. Where strong coupling is important the constant time version of the experiment will suffer additional losses in a manner that depends on the C$^\gamma$:C$^\delta$ chemical shift difference vs coupling ratio.

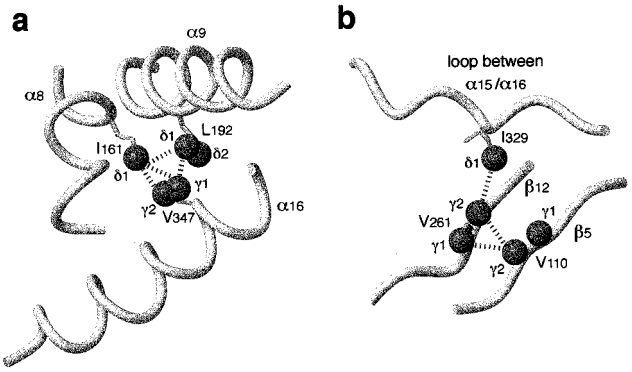

Figure 3. Schematic diagrams showing (a) helices α8, α9, and α16 and (b) strands β5 and β12 and the loop connecting α15/α16 from the X-ray structure of the complex of MBP and β-cyclodextrin[30] indicating how the observed methyl−methyl NOEs define important structural contacts. (a) NOEs between Ile 161, Leu 192, and Val 347 (Figure 2a) and (b) between Val 110, Val 261, and Ile 329 (Figure 2c). The figure was generated with MOLMOL.[39]

The basic experimental scheme described in Figure 1 can, in principle, also be used to record NOEs connecting any pair of proximal protons. In this case the I-BURP2 pulses must be replaced by nonselective inversion pulses. Because the relaxation properties of non-methyl carbons are not nearly as favorable as their methyl counterparts it is especially advisable to record the t_2 evolution period in nonconstant time mode in this case. In addition care must be taken to ensure that spectral widths are properly adjusted so that overlap between folded and unfolded cross-peaks is minimal. In applications discussed in the present paper we focus exclusively on experiments optimized for measuring methyl−methyl correlations.

It is also possible to simultaneously record NOEs between proximal methyls and between neighboring methyl and NH protons by a number of simple modifications to the pulse schemes in parts a and b of Figure 1. In this case the INEPT transfers that sandwich the t_2 evolution period must relay magnetization between NH and ^{15}N in addition to the CH,^{13}C transfers that occur in the sequences of parts a and b of Figure 1. A pulse sequence for recording correlations of the form (ω_{Hi},ω_{Ck},ω_{CHk}) and (ω_{Hi},ω_{Nj},ω_{NHj}), where residues i and k and i and j are close in space, has already been described[29] and the building blocks in that sequence can simply be inserted into the present experiment. Parts c and d in Figure 1 illustrate the necessary modifications, allowing cross-peaks at (ω_{Ci},ω_{Ck},ω_{CHk}) and (ω_{Ci},ω_{Nj},ω_{NHj}) to be recorded.

Figure 2 illustrates F_1−F_3 strips from the 3D ^{13}C,^{13}C methyl NOESY spectrum recorded on MBP in complex with β-cyclodextrin (mixing time of 100 ms) with use of the scheme of Figure 1a. Each slice is taken at the F_2 chemical shift corresponding to the resonance frequency of the carbon indicated at the top of each slice. Our choice of the double constant time version of the experiment in this application was motivated by the fact that MBP is comprised of 370 residues containing 122 Val, Leu, and Ile (C$^{\delta 1}$) methyl groups and issues of cross-peak overlap are therefore very significant. Note that only Val, Leu, and Ile (C$^{\delta 1}$) methyls are protonated in the scheme that we use.[10]

(25) Frankiel, T.; Bauer, C.; Carr, M. D.; Birdsall, B.; Feeney, J. *J. Magn. Reson.* **1990**, *90*, 420−425.
(26) Ikura, M.; Bax, A.; Clore, G. M.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 9020−2022.
(27) Morris, G. A.; Freeman, R. *J. Am. Chem. Soc.* **1979**, *101*, 760−762.
(28) Geen, H.; Freeman, R. *J. Magn. Reson.* **1991**, *93*, 93−141.
(29) Pascal, S. M.; Muhandiram, D. R.; Yamazaki, T.; Forman-Kay, J. D.; Kay, L. E. *J. Magn. Reson. Ser. B* **1993**, *103*, 197−201.

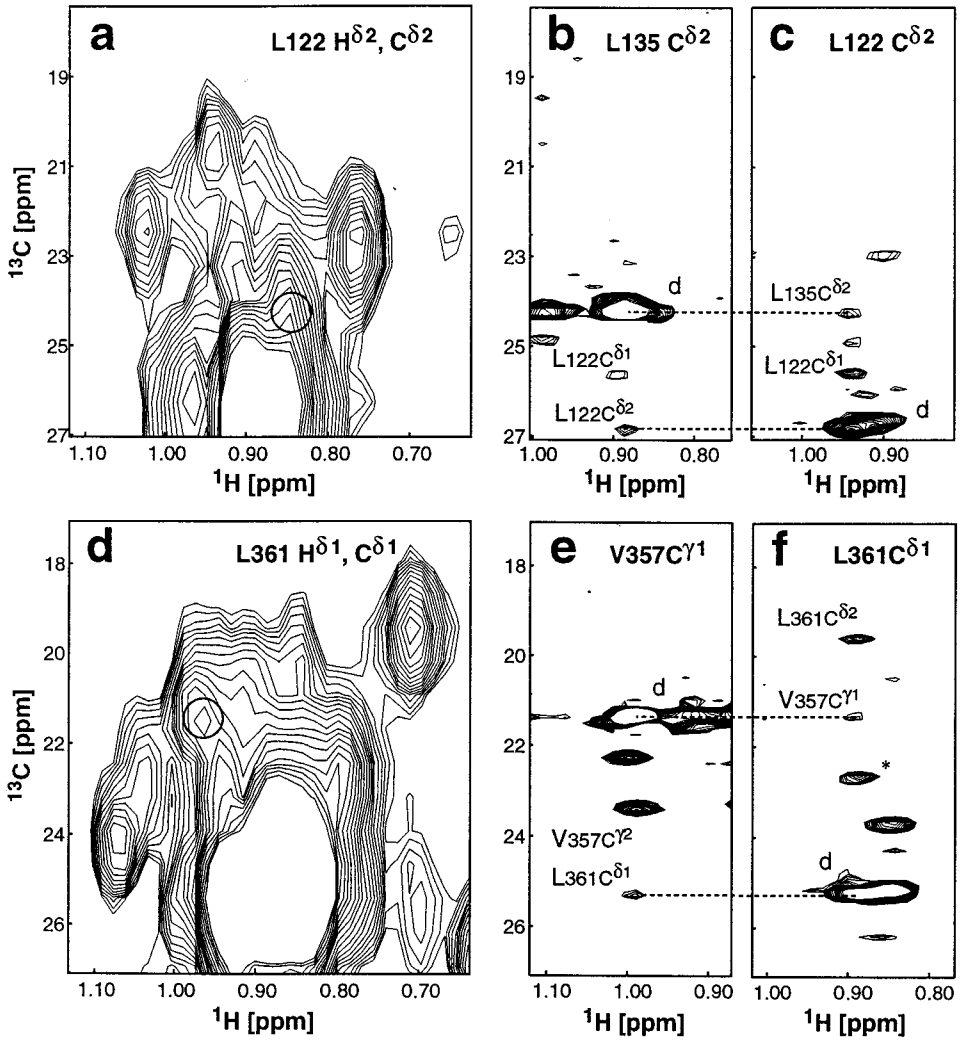

Figure 4. Comparison of slices from the 4D $^{13}C-^{13}C$ NOESY (a, d) and the 3D $^{13}C,^{13}C$ methyl NOESY (b, c, e, f) of the MBP/β-cyclodextrin complex showing NOEs between Leu 122 and Leu 135 (a, b, c) and Leu 361 and Val 357 (d, e, f). The location of the correlation peaks in each of the slices from the 4D data set is indicated with circles.

Moreover, the high quality of spectra obtained with this molecule suggest that a trade of sensitivity for resolution in this case would be worthwhile. In Figure 2a correlations connecting residues Val 347, Ile 161, and Leu 192 (diagonal peaks are marked with "d") are shown along with symmetry-related connectivities, denoted by the horizontal lines in the set of figures. These contacts bring together three helices, α16 (Val 347), α8 (Ile 161), and α9 (Leu 192). In Figure 2b, contacts between Val 183, located in a loop connecting β9 with α9, and Val 343 on α16 are illustrated. Correlations between Val 110, Val 261, and Ile 329, located on β5, β12, and the loop between α15 and α16, respectively, are illustrated in Figure 2c. Finally, NOEs connecting loops joining α4/α5 (Leu 89) and α12/α13 (Leu 285) are indicated in Figure 2d. A large number of additional long-range contacts are shown in Figure 2. It is important to emphasize that the X-ray derived structure[30] was not used in the assignment of any of the 3D NOE data. Over 120 medium and long-range correlations were assigned exclusively on the basis of the high spectral resolution of the data and the symmetry-related connections provided in the spectrum.

The utility of the long-range NOEs obtained in the present experiment for structure determination is illustrated in Figure 3. In Figure 3a the methyl groups involved in the correlations observed in Figure 2a are shown, superimposed on the X-ray structure of the complex.[30] These NOEs provide critical restraints for establishing the correct orientation of helices α8, α9, and α16. As we have discussed in some detail previously, distances between amide protons on proximal helices generally exceed 5−6 Å,[7] and it is therefore often not possible to obtain sufficient restraints to link these elements from backbone NH−NH NOE correlations alone. The situation is even more acute for helices α8 and α9, which are close to perpendicular to each other. The present example provides a powerful illustration of the utility of methyl−methyl correlations for structure determination in general, and in the particular case where only a limited subset of distances are available, such as in the case of highly deuterated molecules. Figure 3b shows the contacts between Val 110, Val 261, and Ile 329 (see Figure 2c), which help orient strands β5 and β12 with respect to the loop between helices α15 and α16.

It is of interest to compare the present 3D data set with the corresponding 4D $^{13}C,^{13}C$ edited NOESY recorded on MBP. It is clear that resolution will be limiting given the short acquisition times employed in the 4D spectrum (4.4 ms in each of the carbon dimensions) and the large number of methyl groups in MBP. This is particularly the case for correlations linking Leu and/or Val residues since the large numbers of these amino acids coupled with their particularly poor dispersion makes assignment

(30) Sharff, A. J.; Rodseth, L. E.; Quiocho, F. A. *Biochemistry* **1993**, *32*, 10553−10559.

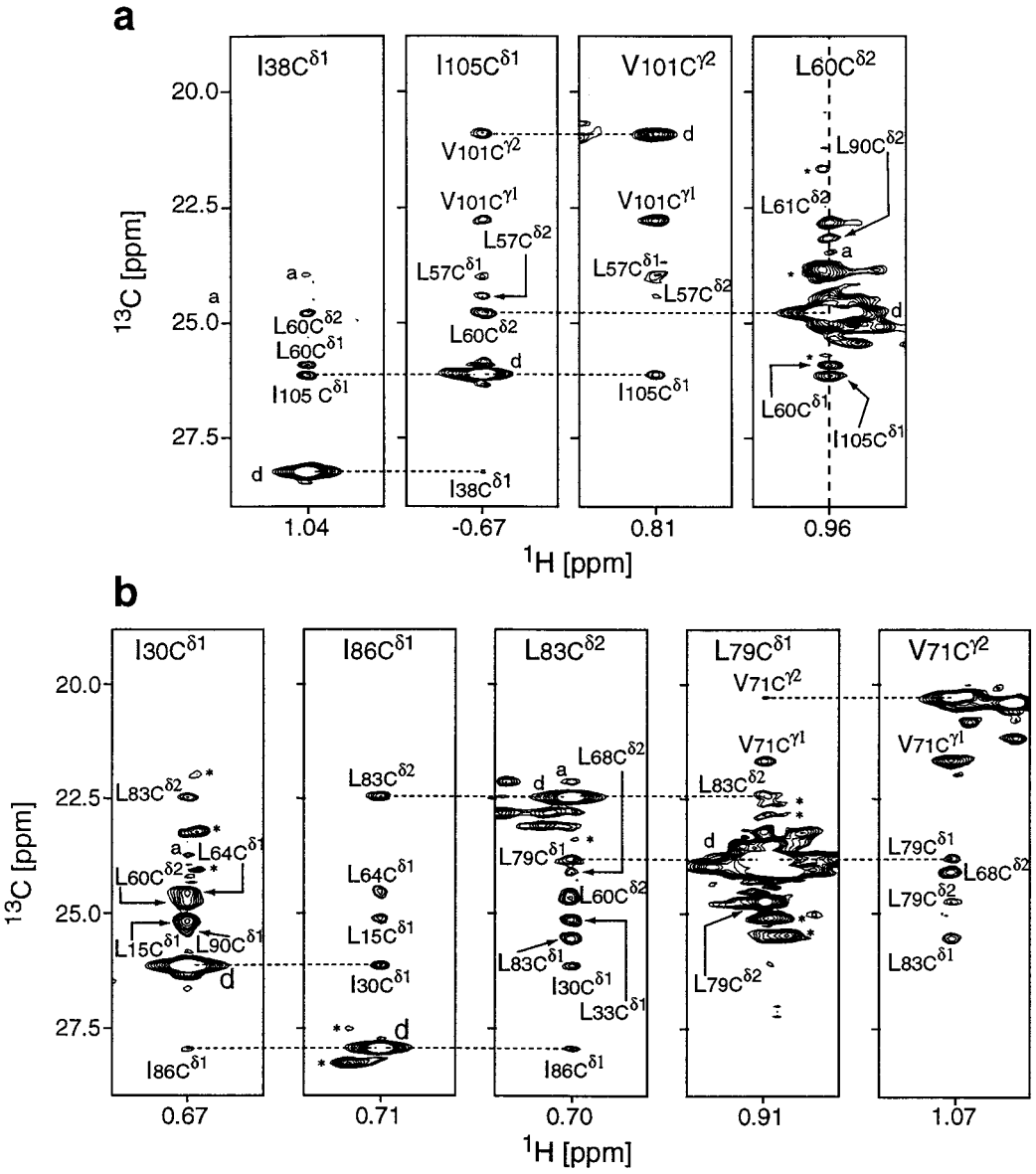

Figure 5. Selected F_1-F_3 strips from the ^{13}C,^{13}C methyl NOESY of methyl-protonated-(Val, Leu, Ile C$^{\delta}$), ^{15}N,^{13}C, ^{2}H-labeled STAT4$_{1-124}$ recorded with the sequence of Figure 1b (mixing time of 175 ms).

of connectivities involving these residues extremely challenging. Parts a and d of Figure 4 illustrate the problems associated with the 4D data set, as described above. In Figure 4a a portion of a 2D correlation map from the 4D spectrum with connectivities arising from the H$^{\delta 2}$ proton of Leu 122 is shown. The high degree of overlap limits the utility of the data. Parts b and c of Figure 4 illustrate correlations between Leu 135 and 122, where the excellent resolution allows the straightforward assignment of the Leu 135, Leu 122 NOE. The location of this NOE in the 4D data set is illustrated by the circle. A similar situation occurs in Figure 4d where the position of the correlation between Leu 361 H$^{\delta 1}$ and Val 357 H$^{\gamma 1}$ is circled. Symmetry related peaks for the Leu 361, Val 357 correlation are observed in parts e and f of Figure 4.

Figure 5 illustrates F_1-F_3 strips from the 3D ^{13}C,^{13}C methyl NOESY (mixing time of 175 ms) recorded on hSTAT4$_{1-124}$ with use of the scheme of Figure 1b with modifications as described in the Materials and Methods section. Because this complex is comprised of a symmetric homodimer there are approximately half the number of methyl groups relative to MBP (59 vs 122). As such, the resolution available in the single constant time version of the experiment was sufficient to allow the unambiguous assignment of close to 50 long-range methyl−methyl contacts. Long-range correlations connecting Ile 38, Leu 60, Val 101, and Ile 105 are shown in Figure 5a. The excellent resolution in F_1 allows separation of NOEs between Ile 38 H$^{\delta 1}$, Leu 60 H$^{\delta 1}$ and Ile 38 H$^{\delta 1}$, Ile 105 H$^{\delta 1}$ as well as the NOEs correlating Leu 60 H$^{\delta 2}$, Leu 60 H$^{\delta 1}$ and Leu 60 H$^{\delta 2}$, Ile 105 H$^{\delta 1}$. In particular, Leu 60 H$^{\delta 2}$, Leu 60 H$^{\delta 1}$ and Leu 60 H$^{\delta 2}$, Ile 105 H$^{\delta 1}$ correlations would be superimposed in spectra acquired with low digital resolution, such as 4D ^{13}C−^{13}C NOESY data sets. In such spectra the connectivity between Leu 60 and Ile 105 would most likely not be assigned, since the correlation between geminal methyls would be expected and assigned to the composite cross-peak. Correlations between Ile 30, Val 71, Leu 79, Leu 83, and Ile 86 are illustrated in Figure 5b. The NOEs between Leu 83 H$^{\delta 2}$ (destination) and the source protons Ile 30 H$^{\delta 1}$, Leu 33 H$^{\delta 1}$, Leu 60 H$^{\delta 2}$, and Leu 79 H$^{\delta 1}$ are particularly noteworthy since all of the carbon shifts of the source methyls lie within 2.3 ppm.

The long-range NOEs depicted in Figure 5 are critical for determining the structural topology of hSTAT4$_{1-124}$, which is

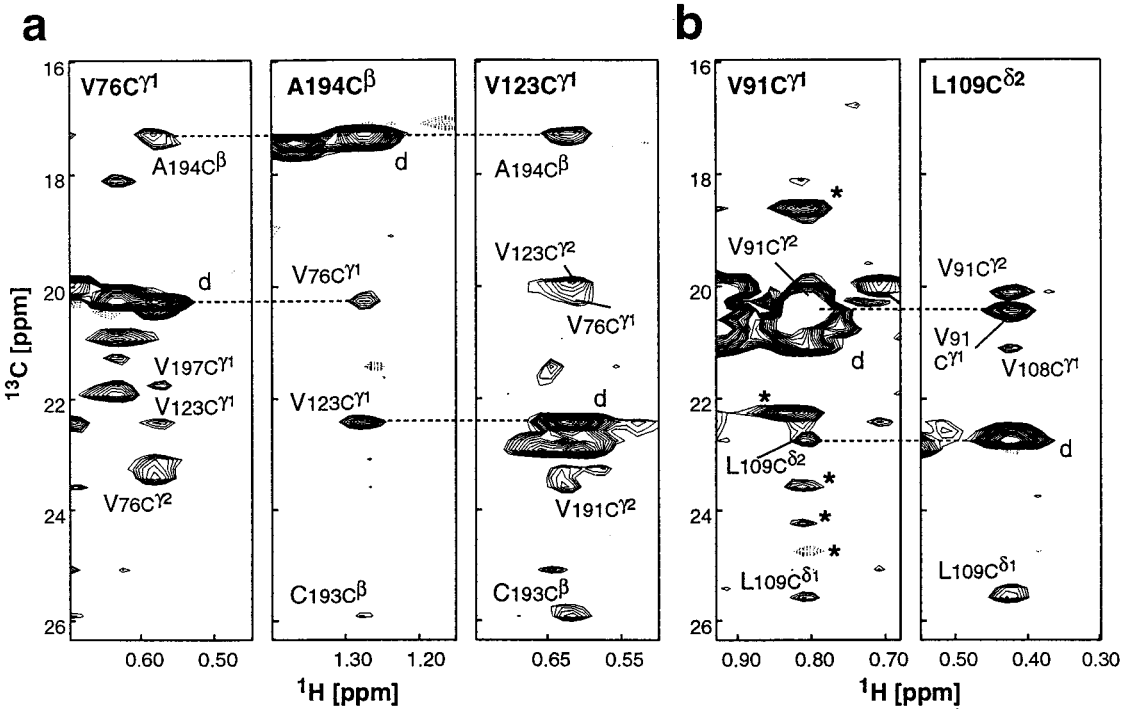

Figure 6. Strips from the $^{13}C,^{13}C$ methyl NOESY of a fully protonated, $^{15}N,^{13}C$-labeled complex of the dNumb PTB domain with peptide (unlabeled) recorded with the scheme of Figure 1b (mixing time of 90 ms).

highly helical.[15,31] The contacts in panel a connect helices α4 (Ile 38), α8 (Ile 105, Val 101), and α6 (Leu 60) while the correlations in panel b link helices α3 (Ile 30), α7 (Ile 86, Leu 83, and Leu 79) and α6 (Val 71). As in the case of MBP, assignment of methyl NOEs was based exclusively on data available from the $^{13}C,^{13}C$ methyl NOESY data set. The 46 long-range methyl−methyl contacts provided in the present data greatly facilitated the calculation of the global fold of the monomer (Sarma and Byrd, unpublished data), which is in agreement with the recently reported X-ray structure of a similar N-terminal domain of murine STAT4.[31]

The utility of the 3D $^{13}C,^{13}C$ methyl NOESY experiment (Figure 1) for protonated molecules has also been established by recording spectra on a protonated, $^{15}N,^{13}C$ dNumb PTB domain in complex with a high affinity binding seven-residue peptide. Although this molecule is of moderate size (160 residues), the overall correlation time established by ^{15}N relaxation experiments[32] on the complex is 13.5 ns. Many of the triple resonance experiments, as well as the 4D NOE-based experiments, have yielded spectra of poor sensitivity, and spectra recorded on this molecule provide, therefore, an excellent test case to establish the efficacy of the present methodology. A comparison of data sets with use of the schemes of parts a and b of Figure 1 indicate that for this molecule the sensitivity gains of the nonconstant time t_2 version of the experiment far outweigh the benefits of improved resolution provided by the double constant time ^{13}C evolution periods. Analysis of the spectra of the PTB domain and of $hSTAT4_{1-124}$ indicates that for applications involving many protein systems typically studied by NMR the resolution provided by the single constant time version of the experiment (Figure 1, parts b and d) will suffice. Figure 6 shows a number of cross sections from a 90 ms 3D $^{13}C,^{13}C$ methyl NOESY spectrum recorded on the PTB domain. In Figure 6a correlations linking Val 76 (helix α1), Val 123 (strand β2), and Ala 194 (helix α3) are illustrated. A significant number of long-range NOEs are observed, including a connectivity between Val 123 ($H^{\gamma 1}$) and Cys 193 (H^{β}). Note that the chemical shift of the C^{β} carbon of Cys 193 (25.85 ppm) is well within the excitation bandwidth of the selective I-BURP2 pulses used during the INEPT transfer schemes. In general, we have found that cross-peaks arising from non-methyl protons do not complicate data analysis. Figure 6b shows symmetry-related correlations between Val 91 ($H^{\gamma 1}$) and Leu 109 ($H^{\delta 2}$) connecting secondary structure elements $\beta 1'$ (Val 91) and α2 (Leu 109). A complete analysis of the $^{13}C,^{13}C$ methyl NOESY spectrum of the protonated PTB domain establishes that high-quality data connecting proximal methyl groups are obtained from this experiment.

In summary, a pulse scheme for correlating methyl groups in proteins has been presented which provides high-quality data for both protonated and methyl-protonated, highly deuterated samples. The methodology is useful for applications to high molecular weight proteins or to proteins where spectral overlap is significant. Distances connecting methyl groups are extremely valuable restraints for obtaining global folds of proteins and initial structures generated from these connectivities can subsequently be used in refinement involving a larger body of NOE information. In this regard the $^{13}C,^{13}C$ methyl NOESY is an important addition to the array of multidimensional NOE experiments that has been developed over the past several years.

(31) Vinkemeir, U.; Moarefi, I.; Darnell, J. E.; Kuriyan, J. *Science* **1998**, *279*, 1048−1052.
(32) Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E. *Biochemistry* **1994**, *33*, 5984−6003.
(33) Patt, S. L. *J. Magn. Reson.* **1992**, *96*, 94−102.
(34) Boyd, J.; Scoffe, N. *J. Magn. Reson.* **1989**, *85*, 406−413.
(35) McCoy, M.; Mueller, L. *J. Am. Chem. Soc.* **1992**, *114*, 2108−2110.
(36) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335−338.
(37) McCoy, M. A.; Mueller, L. *J. Magn. Reson.* **1992**, *98*, 674−679.
(38) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393−399.
(39) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.

Acknowledgment. This research was supported by grants from the National Cancer Institute of Canada (L.E.K.) and the Medical Research Council of Canada (L.E.K.). The research was sponsored in part by the National Cancer Institute, DHHS, under contract with ABL (S.P.S, D.A.H., and R.A.B.). The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The authors thank Professor Kalle Gehring (McGill) for providing the MBP plasmid, Mr. Randall Willis (Hospital for Sick Children) for preparation of the MBP sample, Drs. S. C. Li and Tony Pawson (Mount Sinai Hospital) for the dNumb PTB sample, and Dr. David Waugh and Holly A. Baden (NCI-FCRDC) for providing the hSTAT4$_{1-124}$ clone and for purification of the isotopically labeled protein used in this study. C.Z. is the recipient of a Human Frontiers Science Program Fellowship and K.H.G. was supported by a postdoctoral fellowship from the Helen Hey Whitney Foundation. L.E.K is an International Howard Hughes Research Scholar.

JA981205Z