J|A|C|S
COMMUNICATIONS

Published on Web 03/05/2004

Sequence-Specific Assignments of Methyl Groups in High-Molecular Weight Proteins

Daiwen Yang,*,† Yu Zheng,† Dingjiang Liu,‡ and Daniel F. Wyss‡

Department of Biological Sciences and Department of Chemistry, National University of Singapore, 14 Science Drive 4, Singapore 117543, and Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, New Jersey 07033

Received October 17, 2003; E-mail: dbsydw@nus.edu.sg

Assignments of ^{1}H and ^{13}C resonances in protein side chains are crucial for solution structure determination by NMR. Such assignments are often obtained from HCCH-TOCSY- and H(C)CONH-TOCSY-type experiments[1,2] on the basis of backbone assignments. For proteins larger than 30 kDa, assignments of backbone resonances may be facilitated by deuteration and TROSY.[3] Unfortunately, deuteration reduces the number of protons that provide NOE-based distance constraints for 3D structure determination. Nevertheless, global folds can be determined on the basis of backbone ^{1}HN−^{1}HN NOEs and dipolar couplings. To improve the precision and accuracy of the global folds determined from highly deuterated samples, selective reintroduction of methyl protons into methyl-containing residues is critical.[4] Assignment of selectively protonated methyl groups can be achieved from TOCSY-based experiments or the TROSY versions of these experiments.[2b−d] However, because of the low efficiencies of TOCSY transfer and the subsequent magnetization transfer from ^{13}C$^{\alpha}$ to ^{1}HN, the H(C)CONH-TOCSY-type experiments have very low sensitivities in the case of proteins with overall correlation times larger than 20 ns. Recently, a 3D (H)C(CA)-COSY experiment has been proposed to improve the magnetization transfer efficiency, and it is a factor ∼1.5 times more sensitive than a similar TOCSY-based approach for a maltose binding protein at 5 °C.[5]

We propose here a novel 3D multiple-quantum (MQ) (H)CC$_m$H$_m$-TOCSY experiment for assignments of ^{1}H and ^{13}C resonances of methyl groups using uniformly ^{13}C-labeled samples. The new 3D MQ-(H)CC$_m$H$_m$-TOCSY experiment correlates chemical shifts of aliphatic carbon nuclei of amino acid side chains with those of the methyl ^{13}C$_m$ and ^{1}H$_m$ nuclei in the same residue in the protein sequence. On the basis of prior assignments of ^{13}C$_{\alpha}$ and ^{13}C$_{\beta}$, sequence-specific assignment of methyl resonances can be obtained. Figure 1 shows the pulse sequence of the 3D MQ-(H)CC$_m$H$_m$-TOCSY experiment. The magnetization transfer is shown schematically as follows:

$$^1\mathrm{H} \xrightarrow{^1J_{CH}} \mathrm{MQ}(t_1) \xrightarrow{^1J_{CH}} {}^{13}\mathrm{C} \xrightarrow{\mathrm{TOCSY}} {}^{13}\mathrm{C}_m(\mathrm{CT}\ t_2) \xrightarrow{^1J_{CH}} {}^1\mathrm{H}_m(t_3)$$

The decay rates of the MQ coherences (H$_x$C$_y$) are normally significantly smaller than those of the single-quantum (SQ) coherences.[6] Thus, this experiment is more sensitive than its SQ version. During the first part of the constant-time period (point *b* to point *c*), proton decoupling is applied to maintain the slow decays of the transverse ^{13}C$_m$ magnetizations, enhancing experimental sensitivity. In the absence of proton decoupling, proton spin flip−flop rates dominate decays of the two inner components of ^{13}C$_m$ quartets for fully protonated proteins and significantly elevate the apparent ^{13}C$_m$ decays.[7]

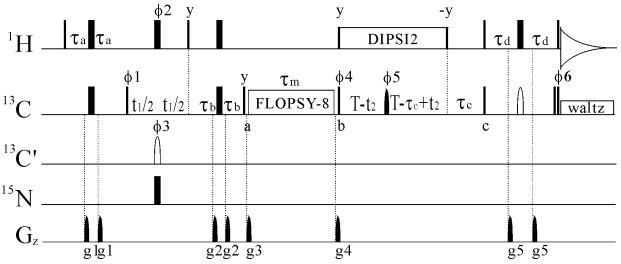

Figure 1. Pulse sequence for the MQ-(H)CC$_m$H$_m$-TOCSY experiment. All narrow (wide) rectangular pulses have flip angles of 90° (180°). The ^{1}H carrier is set at 4.7 ppm while the ^{13}C carrier is centered at 41 ppm, until immediately prior to the ^{13}C pulse of phase ϕ_4 at which time the carrier is jumped to 17 ppm. All ^{1}H pulses are applied with a 23 kHz field; ^{1}H DIPSI2-decoupling elements make use of a 6.25 kHz field. All ^{13}C rectangular pulses employ a 16.8 kHz field, and the ^{13}C shaped pulses have REBURP profiles. The first 180° shaped (filled) ^{13}C pulse has a duration of 400 μs and is phase modulated by 24 ppm, while the second one (empty) has a duration of 1.5 ms. The ^{13}C spin-lock field strength for FLOPSY is 7 kHz. A decoupling power of 1.25 kHz is used during acquisition. The 180° pulse on C′ has a SEDUCE profile with a duration of 250 μs (center of excitation 176 ppm). The delays used are: $\tau_a = 1.4$ ms, $\tau_b = 1.1$ ms, $\tau_c = 1.5$ ms, $\tau_d = 1.6$ ms, $\tau_m = 17$ ms, which is suitable for proteins with overall correlation times ranging from 20 to 30 ns as shown by numerical simulations; $T = 14$ ms. The phase cycling employed is: $\phi_1 = 4(x), 4(-x)$; $\phi_2 = x, y, -x, -y$; $\phi_3 = 2(x), 2(-x)$; $\phi_4 = y$; $\phi_5 = 2(x), 2(y), 2(-x), 2(-y)$; $\phi_6 = 4(x), 4(-x)$; rec $= x, -x, -x, x, -x, x, x, -x$. The duration and strengths of the sine-shaped gradients are: $g1$ = (0.5 ms, 20 G/cm); $g2$ = (0.3 ms, 25 G/cm); $g3$ = (1 ms, 25 G/cm); $g4$ = (1 ms, 20 G/cm); $g5$ = (0.5 ms, 20 G/cm). Quadrature detection in F1 and F2 are achieved by States-TPPI of ϕ_1 and ϕ_4, respectively.

We have applied the pulse sequence shown in Figure 1 to a ^{13}C,^{15}N-labeled sample of acyl carrier protein synthase (AcpS) which consists of three subunits with a total molecular weight of 42 kDa.[8] As established from ^{15}N relaxation data, AcpS has an overall rotational time of 26 ns at 25 °C (equivalent to a protein on the order of 60 kDa at 37 °C). Despite its large overall correlation time, all of the aliphatic ^{13}C resonances were observed for most residues having methyl groups in the MQ-(H)CC$_m$H$_m$−TOCSY experiment. Figure 2 shows a number of F1−F3 slices taken from a 3D MQ-(H)CC$_m$H$_m$-TOCSY spectrum. If the (^{13}C$_{\alpha}$, ^{13}C$_{\beta}$) chemical shifts of residues containing methyl groups are not degenerate with each other in a given protein, sequence-specific assignment of methyl resonances can be obtained from the MQ-(H)CC$_m$H$_m$-TOCSY spectrum on the basis of the assignment of (^{13}C$_{\alpha}$, ^{13}C$_{\beta}$). For example, according to the (^{13}C$_{\alpha}$, ^{13}C$_{\beta}$) chemical shifts and spectral pattern shown in slice e and the prior sequential assignment that was obtained using a uniformly ^{2}H,^{13}C,^{15}N-labeled protein,[8] signals in this slice were assigned to attribute to I96γ_2. ^{13}C$_{\gamma 1}$ and ^{13}C$_{\delta 1}$ from the same slice can then also be assigned. The assignment of ^{13}C$_{\delta 1}$ can be further confirmed from the slice taken at F2 frequency of ^{13}C$_{\delta 1}$ (slice f). Sometimes, ^{13}C$_{\alpha}$ or ^{13}C$_{\beta}$ resonances are not observable in the slice taken from ^{13}C$_{\delta 1}$ of Ile, but observable

† National University of Singapore.
‡ Schering-Plough Research Institute.

10.1021/ja039102q CCC: $27.50

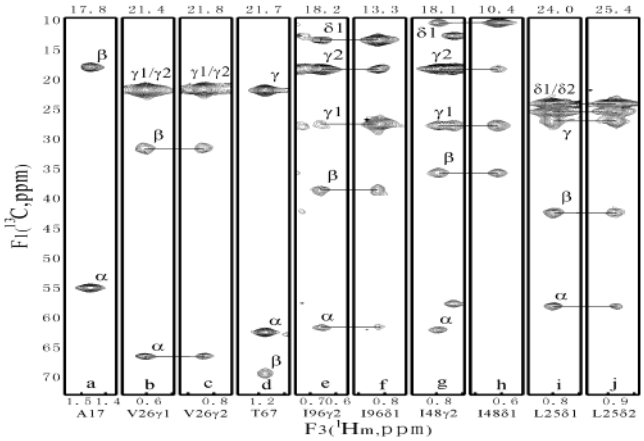

Figure 2. Representative slices from the MQ-(H)CC$_m$H$_m$-TOCSY spectrum used for methyl assignments. Each F1(^{13}C)−F3(^{1}H$_m$) slice is labeled with the identity of the methyl-containing residue, and the F2(^{13}C$_m$) frequency in ppm is indicated at the top of each slice. The experiment was recorded with the 42 kDa uniformly ^{13}C,^{15}N-labeled AcpS homotrimer in ^{1}H$_2$O:^{2}H$_2$O (95:5) solution (protein concentration 0.4 mM in the trimer, pH 7.5, 25 °C) on a Bruker Avance 500 MHz spectrometer equipped with a CryoProbe. 64(t_1) × 70(t_2) × 512(t_3) complex points were collected, giving t_{1max} = 7.9 ms, t_{2max} = 27.4 ms and t_{3max} = 64 ms. An interscan delay of 1 s with 8 scans per increment was used, resulting in a total experimental time of 47 h. The ^{13}C and ^{13}C$_m$ time domains were doubled by forward−backward and mirror-image linear prediction respectively, prior to the application of cosine-squared window functions.

in the slice of ^{13}C$_{\gamma 2}$. In this case, assignment of ^{13}C$_{\delta 1}$ and ^{1}H$_{\delta 1}$ can be done on the basis of the assignment of ^{13}C$_{\gamma 2}$ and the matches of ^{13}C$_{\gamma 1}$, ^{13}C$_{\delta 1}$, and ^{13}C$_{\gamma 2}$ resonances between the two slices as shown in slices g and h. Similarly, the spectral information of two ^{13}C$_\delta$ in Leu or two ^{13}C$_\gamma$ in Val residues can be complementary to each other.

Using the strategy described above, we obtained sequence-specific assignments for 63 out of 67 methyl groups in AcpS. The (^{13}C$_\alpha$, ^{13}C$_\beta$) chemical shifts of I10 and I29 are identical and thus cannot be assigned uniquely. For large monomeric proteins, one would expect that degeneracy in (^{13}C$_\alpha$, ^{13}C$_\beta$) chemical shifts may hinder the application of the method proposed here. In practice, the relatively good dispersion of (^{13}C$_\alpha$, ^{13}C$_\beta$) chemical shifts among the same types of residue allows one to assign most of the methyl resonances, according to our survey as shown in Table 1 (in Supporting Information). Although the (^{13}C$_\alpha$, ^{13}C$_\beta$) chemical shifts of L18 and L25 are degenerate within a threshold of 0.3 ppm, L18 and L25 can be distinguished from each other from the differences of their (^{13}C$_\alpha$, ^{13}C$_\beta$) chemical shifts (0.15 ppm). Obviously, a larger number of methyl groups can be assigned with increasing spectral resolution in the F1 dimension. It is interesting to note that each type of amino acid displays a specific spectral pattern, and thus, (^{13}C$_\alpha$, ^{13}C$_\beta$) chemical shift degeneracy among different types of amino acids is not an issue for methyl assignments.

L9 and L51 displayed very weak ^{13}C$_\alpha$ and ^{13}C$_\beta$ resonances in the MQ-(H)CC$_m$H$_m$-TOCSY experiment. This is attributed to the strong scalar coupling interaction between ^{13}C$_\delta$ and ^{13}C$_\gamma$ spins in Leu residues. To achieve a high resolution in the F2 dimension, a constant-time acquisition mode was used to remove ^{13}C−^{13}C scalar coupling effects. For Leu, however, the strong scalar coupling interaction can be destructive to the refocus of ^{13}C$_\delta$ magnetization during the constant-time period which leads to a significant loss of sensitivity. Although a nonconstant-time version of the experiment that is similar to the HCCH-TOCSY should give better sensitivity, the resolution may not be sufficient to uniquely assign methyl resonances in the case where ^{13}C−^{1}H HSQC cross-peaks are not unique within a grid of 0.3 ppm (^{13}C) × 0.02 ppm (^{1}H). Resolution in the F2 dimension is critical for methyl assignments of large and medium-sized proteins; e.g., only 62% of methyl groups in MBP are not degenerate within the 0.3 ppm × 0.02 ppm grid.

A comparison of the MQ-(H)CC$_m$H$_m$-TOCSY experiment with the HCCH-TOCSY experiment (Figure 2 in Supporting Information) shows that the MQ-(H)CC$_m$H$_m$-TOCSY is more sensitive for most cross-peaks due to the gains from MQ line narrowing in the F$_1$ dimension and from the slow decay of in-phase methyl ^{13}C magnetizations during the constant-time t_2 period. Most importantly, however, most methyl resonances cannot be assigned using the HCCH-TOCSY spectrum due to poor resolution in both the F$_1$ and F$_2$ dimensions and poor dispersion of (^{1}H$_\alpha$, ^{13}C$_\alpha$) and (^{1}H$_\beta$, ^{13}C$_\beta$) chemical shifts. Compared to the C(CO)NH-TOCSY experiments, the experiment proposed here is much more sensitive (∼7 times for ^{13}C-labeled AcpS). The 4D HCCH-NOESY experiment[9] is also less sensitive than the MQ-(H)CC$_m$H$_m$-TOCSY experiment due to the inherently long NOE mixing time and the additional dimension involved. Similar to the HCCH-TOCSY spectrum, most methyl groups may not be assignable because of the low spectral resolution in the 4D and the poorly dispersed chemical shifts of (^{1}H, ^{13}C) spin pairs in large proteins. Compared to TOCSY-based methods established previously,[2c,d] the method proposed here is more sensitive and efficient since methyl assignments require only a single experiment rather than two or three 3D experiments.

The experiment proposed here aims for only ^{13}C-labeled proteins that can be produced more easily than ^{2}H,^{13}C,^{1}H$_m$-labeled proteins. If a ^{2}H,^{13}C,^{1}H$_m$-labeled protein sample is available, one may use an alternative scheme (CC$_m$H$_m$-TOCSY) as shown in Figure 3 (Supporting Information). This experiment can be more sensitive than the MQ-(H)CC$_m$H$_m$-TOCSY experiment when ^{13}CD T_1 can be effectively reduced by paramagnetic relaxation agents, especially for very large proteins (>50 kDa).

Acknowledgment. This research was supported by a grant from the Biomedical Research Council (BMRC) of the Agency for Science, Technology and Research (A*STAR) of Singapore.

Supporting Information Available: One table listing the percentage of methyl-containing residues in which the difference between two ^{13}C$_\alpha$ or ^{13}C$_\beta$ is equal to or larger than 0.3 ppm, one figure showing ^{13}C−^{1}H HSQC with methyl assignments, one figure showing the S/N ratios in two spectra, and one for the CC$_m$H$_m$-TOCSY pulse scheme. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Fesik, S. W.; Eaton, H. L.; Olejniczak, E. T.; Zuiderweg, E. R. P.; McIntosh, L. P.; Dahlquist, F. W. *J. Am. Chem. Soc.* **1990**, *112*, 886.
(2) (a) Montelione, G. T.; Lyons, B. A.; Emerson, S. D.; Tashiro, M. *J. Am. Chem. Soc.* **1992**, *114*, 10974−10975. (b) Gardner, K. H.; Konrat, R.; Kay, L. E. *J. Biomol. NMR* **1996**, *8*, 351−356. (c) Hilty, C.; Fernandez, C.; Wider, G.; Wuthrich, K. *J. Biomol. NMR* **2002**, *23*, 289−301. (d) Gardner, K. H.; Zhang, X. C.; Gehring, K.; Kay, L. E. *J. Am. Chem. Soc.* **1998**, *120*, 11738−11748.
(3) (a) Pervushin, K.; Riek, R.; Wider, G.; Wuthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12366−12371. (b) Tugarinov, V.; Muhandiram, R.; Ayed, A.; Kay, L. E. *J. Am. Chem. Soc.* **2002**, *124*, 10025−10035. (c) Yang, D.; Kay, L. E. *J. Am. Chem. Soc.* **1999**, *121*, 2571−2575.
(4) Rosen, M. K.; Gardner, K. H.; Willis, R. C.; Parris, W. E.; Pawson, T.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 627−636.
(5) Tugarinov, V.; Kay, L. E. *J. Am. Chem. Soc.* **2003**, *125*, 5701−5706.
(6) (a) Grzesiek, S.; Kuboniwa, H.; Hinck, A. P.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 5312−5315. (b) Gschwind, R. M.; Gemmecker, G.; Kessler, H. *J. Biomol. NMR* **1998**, *11*, 191−198. (c) Shang, Z.; Swapna, G. V. T.; Rios, C. B.; Montelione, G. T. *J. Am. Chem. Soc.* **1997**, *119*, 9274−9278.
(7) Liu, W. D.; Zheng, Y.; Cistola, D. P.; Yang, D. W. *J. Biomol. NMR* **2003**, *27*, *351−364*.
(8) Liu, D. J.; Black, T.; Macinga, D. R.; Palermo, R.; Wyss, D. F. *J. Biomol. NMR* **2002**, *24*, 273−274.
(9) Fischer, M. W. F.; Zeng, L.; Zuiderweg, E. R. P. *J. Am. Chem. Soc.* **1996**, *118*, 12457−12458.

JA039102Q