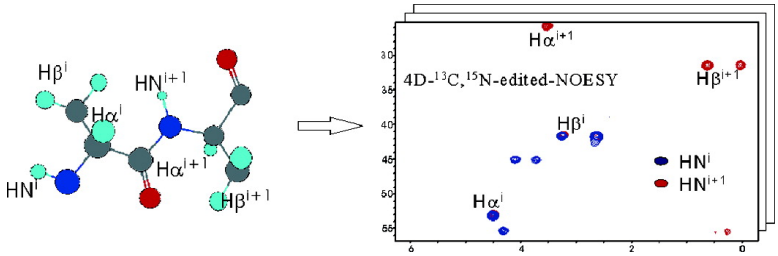
Communication

A General Strategy for the Assignment of Aliphatic Side-Chain Resonances of Uniformly ^{13}C,^{15}N-Labeled Large Proteins

Yingqi Xu, Zhi Lin, Chien Ho, and Daiwen Yang

J. Am. Chem. Soc., **2005**, 127 (34), 11920-11921• DOI: 10.1021/ja053539b • Publication Date (Web): 04 August 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 08/04/2005

A General Strategy for the Assignment of Aliphatic Side-Chain Resonances of Uniformly ^{13}C,^{15}N-Labeled Large Proteins

Yingqi Xu,[†] Zhi Lin,[†] Chien Ho,[‡] and Daiwen Yang*,[†]

Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, and Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Received May 31, 2005; E-mail: dbsydw@nus.edu.sg

For proteins smaller than 25 kDa, backbone and side-chain assignments can be obtained using uniformly ^{13}C,^{15}N-labeled proteins with triple resonance experiments. Thus, structure determination by NMR is very suitable for this type of small proteins.[1] The same methods may fail, however, when applied to proteins larger than 30 kDa due to increased transverse relaxation rates. With the introduction of deuteration and TROSY techniques, backbone and $^{13}C^{\beta}$ resonances can be assigned for proteins up to 100 kDa.[2,3] Unfortunately, deuteration also significantly reduces structural constraints derived from NOEs from aliphatic and aromatic protons. Although it is possible to build structural models according to NOEs from only amide protons and residual dipolar couplings in partially ordered medium, the protein structure always suffers from low precision.[4] To improve structure precision, one can selectively reintroduce methyl protons into otherwise deuterated samples since methyl groups are often involved in hydrophobic cores and provide long-range distance constraints.[5] There have been several successful applications of the selective labeling strategy to large proteins.[6] However, preparation of deuterated and methyl-protonated samples is always costly and time-consuming and may not be suitable for every protein. To further improve structure resolution, it is necessary to constrain side chains of all or most residues using NOEs among protons located at side chains. This implies that complete or partial protonation at most side chains is inevitable.

For fully protonated large proteins, our group has recently introduced a novel experiment (MQ-(H)CC$_m$H$_m$-TOCSY)[7] and a strategy to assign side-chain resonances of methyl-containing residues.[8] The method has been successfully applied to a 42 kDa AcpS trimer and a 65 kDa chain-selectively labeled hemoglobin. So far, no method can be applied to assign side-chain resonances of residues that contain no methyl groups.

In this work, we present a general strategy to assign side-chain resonances of all residues in uniformly ^{13}C,^{15}N-labeled large proteins, which makes use of 4D ^{13}C,^{15}N-edited NOESY and prior assignment of backbone. Although most triple resonance experiments involving both ^{13}C and ^{15}N spins have very poor sensitivity for protonated large proteins, NOESY experiments are still sensitive enough to provide through-space correlations between spins separated by 4.5 Å or less (5.5 Å for methyl groups). Statistics on interatomic distances[9] (Table S1 of the Supporting Information) indicate that nearly all intraresidue $H^N_i-H^{\alpha}_i$, $H^N_i-H^{\beta}_i$ and sequential $H^N_i-H^{\alpha}_{i-1}$, $H^N_i-H^{\beta}_{i-1}$ NOEs can be observed, where the subscript i is the residue number. In a ^{13}C,^{15}N-edited 4D NOESY spectrum, each amide correlates with a number of CH_n groups at positions [$\omega(H^N_i)$, $\omega(N_i)$, $\omega(C^k_j)$, $\omega(H^k_j)$] where ω is the chemical shift; k is the kth carbon/ hydrogen of residue j. H^{α} and H^{β} can be assigned from intraresidue or sequential NOEs, provided that these NOEs are possibly differentiated from other interresidue NOEs on the basis of prior assignment of H^N, N, C^{α}, and C^{β} spins. Otherwise, ambiguities in assignment can be resolved using both intraresidue and sequential NH−CH NOE correlations (e.g., [$\omega(H^N_i)$, $\omega(N_i)$, $\omega(C^{\alpha}_i)$, $\omega(H^{\alpha}_i)$] and [$\omega(H^N_{i+1})$, $\omega(N_{i+1})$, $\omega(C^{\alpha}_i)$, $\omega(H^{\alpha}_i)$]). If the ambiguity in assignment cannot be resolved due to a lack of sequential or intraresidue NOEs, an MQ-(H)CCH-TOCSY experiment can be applied to confirm the assignment. Assignments of protons at γ, δ, and ϵ positions are much more challenging using the 4D ^{13}C,^{15}N-edited NOESY since the chemical shifts of carbon spins at these positions are not available. According to the statistics on proton−proton distances, many H^{γ}s and some H^{δ}s give rise to both intraresidue and sequential NH−CH NOEs and thus can be assigned from the NOESY spectrum. The remaining unassigned spins can be assigned using the MQ-(H)CCH-TOCSY and 4D NOESY experiments.

The strategy mentioned above was tested on a cell adhesion protein (DdCAD-1, 214 residues, ∼0.8 mM protein on a 500 MHz NMR) whose backbone and side-chain resonances have been assigned previously using traditional methods.[10] It was applied to the assignment of aliphatic side chains in the uniformly ^{13}C-labeled β-chain of human normal adult hemoglobin in the carbonmonoxy form (rHbCO A, ∼65 kDa with two identical α-chains and two identical β-chains). Its backbone, most methyl groups, and side-chain carbons in methyl-containing residues have been assigned previously.[8] The procedure used for the assignment of side-chain protons and carbons is described below.

First, identify peaks whose chemical shifts match the shifts [$\omega(H^N_i)$, $\omega(N_i)$, $\omega(C^{\alpha}_i)$] for residue i on the N_i-H_i plane in the 4D NOESY. If only one NOE peak matches, the aliphatic proton shift of this peak is presumably assigned as the chemical shift of H^{α}_i. Similarly, by substituting $\omega(C^{\alpha}_i)$ for $\omega(C^{\alpha}_{i-1})$, $\omega(C^{\beta}_i)$, and $\omega(C^{\beta}_{i-1})$, H^{α}_{i-1}, H^{β}_i, and H^{β}_{i-1} can be presumably assigned, provided that unique matches also exist (for CH_2 groups, two peaks with identical carbon shifts are also regarded as a unique match). Since degeneracy of (H^N, N, C) spin triplets occurs in a much lower chance than that of (H^N, N) spin pairs, most H^{α} and H^{β} could be presumably assigned with only intraresidue or sequential NOE (Table 1, columns A and B). Second, if the assignment obtained from intraresidue NOEs is consistent with that obtained from sequential NOEs, the assignment is confirmed (Figure 1a,b). Third, for an unconfirmed assignment obtained from the N_i-H_i plane, if its [$\omega(C_j)$, $\omega(H_j)$] shifts match the shifts of one of the peaks on the $N_{i+1}-H_{i+1}$ or $N_{i-1}-H_{i-1}$ plane, the assignment is also confirmed (Figure 1c,d). When no assignment is obtained from step 1, due to ambiguities, the N_i-H_i and $N_{i+1}-H_{i+1}$ planes can be directly compared to resolve the ambiguities. For DdCAD-1, we found that all assignments obtained with both intraresidue and sequential NOEs (Table 1, column C) were correct, while three of the unconfirmed assignments (Table 1, column D) were incorrect.

Fourth, the same strategy can be used for the assignment of other side-chain resonances. Although exact chemical shifts of C^{γ} and

[†] National University of Singapore.
[‡] Carnegie Mellon University.

10.1021/ja053539b CCC: $30.25

Table 1. Summary of Assignments of DdCAD-1 and rHbCO A[a]

	A	B	C	D	E
		DdCAD-1[b]			
$C^{\alpha}H^{\alpha}_n$ (212)	161	168	170	32	206/5/1
$C^{\beta}H^{\beta}_n$ (199)	172	140	145	44	195/2/2
$C^{\gamma}H^{\gamma}_n$ (72)[c]	38	31	28	25	66/5/1
$C^{\delta}H^{\delta}_n$ (32)[c]	13	8	7	10	29/1/2
CH_3 (101)[d]	37	35	76	9	100/0/1
		β-chain of rHbCO A[b]			
$C^{\alpha}H^{\alpha}_n$ (146)	114	73	90	44	132/12/2
$C^{\beta}H^{\beta}_n$ (133)	97	67	74	31	90/26/17
$C^{\gamma}H^{\gamma}_n$ (51)[c]	23	4	15	12	32/12/7
$C^{\delta}H^{\delta}_n$ (21)[c]	9	6	5	8	10/9/2
CH_3 (80)[d]	19	16	40	4	78/2/0

[a] (A) Assigned with intraresidue NOE; (B) assigned with sequential NOE; (C) assigned with both intraresidue and sequential NOEs; (D) unconfirmed (only intraresidue NOE or sequential NOE was observed); (E) the final assigned/tentatively assigned/ unassigned CH_n groups using both the NOESY and CCH-TOCSY spectra. [b] Total number of CH_n groups is indicated in brackets. [c] Excluding methyls. [d] Excluding Ala.

C^{δ} are not known, their empirical ranges can be used to locate possible peaks (Figure 2a,b). Because of the obvious problem of chemical shift degeneracy as well as usually longer distances to amide protons, less number of $C^{\gamma}H^{\gamma}_n$ and $C^{\delta}H^{\delta}_n$ groups can be assigned with the 4D ^{13}C,^{15}N-edited NOESY alone (Table 1).

Finally, an 3D MQ-(H)CCH-TOCSY spectrum is used in combination with the 4D NOESY to assign the unconfirmed and remaining resonances (Figure 2c−g). For DdCAD-1, nearly all correlations can be observed in the TOCSY and result in aliphatic side-chain assignment completeness of ∼96% (the ratio of the assigned to total aliphatic CH_n groups) (Table 1, column E), which is comparable with that obtained from traditional methods.[10] For Hb, many peaks involving H^{α} or H^{β} cannot be observed in the TOCSY spectrum; in contrast, the peaks involving H^{γ}, H^{δ}, or methyl protons are usually observable due to higher mobility of these spins. Sixteen methyl groups that were ambiguously assigned previously due to degenerate [C^{α}, C^{β}] spin pairs were completely assigned here using the intraresidue CH_3−NH NOEs observed in the 4D

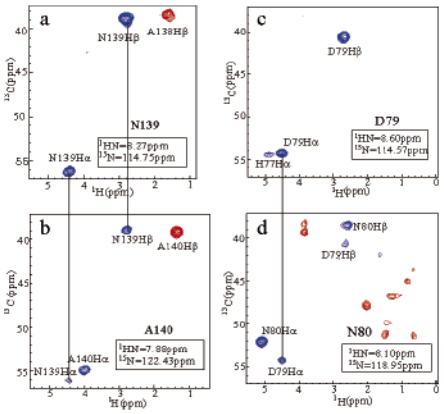


Figure 1. Representative N_k−H_k/F_1(^{1}H)−F_2(^{13}C) planes from the 4D ^{13}C,^{15}N-edited NOESY. Each plane is labeled with its ^{15}N and ^1HN chemical shifts and the corresponding amino acid. All red peaks were aliased by 20 ppm in the ^{13}C dimension. The unlabeled peaks in d are from the neighboring planes. The experiment was recorded with the ^{13}C,^{15}N-labeled β-chains complexed with unlabeled α-chains of Hb in ^1H_2O:^2H_2O (95:5) solution (∼2 × 0.5 mM in the β-chain, pH 7, 30 °C) on a Bruker Avance 500 MHz spectrometer equipped with a CryoProbe; 18(t_1) × 17(t_2) × 18(t_3) × 512(t_4) complex points were collected with a mixing time of 50 ms, giving t_{1max} = 4.8 ms, t_{2max} = 5.6 ms, t_{3max} = 6.4 ms, and t_{4max} = 64 ms. An interscan delay of 1 s with 8 scans per increment was used, resulting in a total experimental time of 112 h. The indirect domains were doubled by forward−backward linear prediction prior to the application of cosine squared window functions.

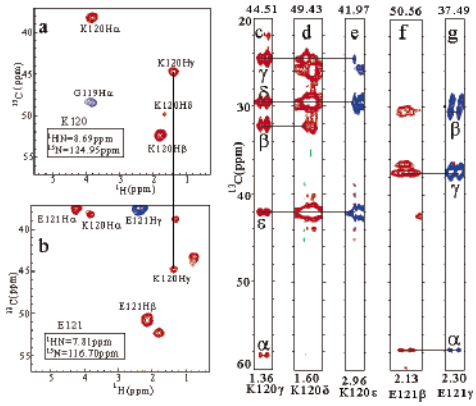


Figure 2. Assignment of $C^{\gamma}H^{\gamma}_n$ and $C^{\delta}H^{\delta}_n$ using the 4D ^{13}C,^{15}N-edited NOESY (a and b) and CCH-TOCSY (c−g) spectra. Red peaks in slices a and b are aliased by 20 ppm in the ^{13}C dimension. The F2 frequencies (ppm) for slices c−g are indicated on the top of each slice. The CCH-TOCSY data comprising 105 × 35 × 640 complex points with spectral widths of 12 007, 4024, and 12 007 Hz in F1, F2, and F3 dimensions were collected on an 800 MHz NMR spectrometer with a triple resonance probe using a TOCSY mixing time of 14 ms, an interscan delay of 1.2 s and 8 scans per increment (41 h).

data (Figure S1 of the Supporting Information). After these steps, ∼80% side-chain spins were assigned for Hb (Table 1, column E). Most unassigned spins lack NH−CH NOEs. More assignments can be obtained using a 4D ^{13}C-edited NOESY experiment, provided the unassigned spins give rise to intraresidue CH−CH NOEs.

In conclusion, most aliphatic side-chain resonances of large proteins can be assigned reliably with 4D ^{13}C,^{15}N-edited NOESY and MQ-(H)CCH-TOCSY experiments, which will make complete use of currently available backbone assignments and provide much more distance constraints for accurate structure determination of large proteins. Although the strategy based on NOESY and TOCSY has been used for peptides and small proteins for many years,[11] it is demonstrated here, for the first time, that a similar strategy can be applied to large proteins up to ∼65 kDa.

Acknowledgment. This research was supported by a Young Investigator Award from the Biomedical Research Council (BMRC), and Agency for Science, Technology and Research, A*Star of Singapore. The authors thank Prof. Lewis E. Kay for helpful discussions.

Supporting Information Available: One table lists statistics on proton−proton distances and one figure shows the assignment of methyl groups. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Bax, A. *Curr. Opin. Struct. Biol.* **1994**, *4*, 738.
(2) Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *J. Am. Chem. Soc.* **2000**, *122*, 7543.
(3) Yang, D. W.; Kay, L. E. *J. Am. Chem. Soc.* **1999**, *121*, 2571.
(4) Giesen, A. W.; Homans, S. W.; Brown, J. M. *J. Biomol. NMR* **2003**, *25*, 63.
(5) Goto, N. K.; Gardner, K. H.; Mueller, G. A.; Willis, R. C.; Kay, L. E. *J. Biomol. NMR* **1999**, *13*, 369.
(6) Tugarinov, V.; Choy, W.-Y.; Orekhov, V. Y.; Kay, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 622.
(7) Yang, D. W.; Zheng, Y.; Liu, D. J.; Wyss, D. F. *J. Am. Chem. Soc.* **2004**, *126*, 3710.
(8) Zheng, Y.; Giovannelli, J. L.; Ho, N. T.; Ho, C.; Yang, D. W. *J. Biomol. NMR* **2004**, *30*, 423.
(9) Zheng, Y.; Yang, D. W. *Bioinformatics* **2005**, *21*, 2925.
(10) Lin, Z.; Huang, H. B.; Siu, C.-H.; Yang, D. W. *J. Biomol. NMR* **2004**, *30*, 375.
(11) Wuthrich, K. *NMR of Proteins and Nucleic Acids*, Wiley: New York, 1986.

JA053539B