Use of $^1H_N-^1H_N$ NOEs To Determine Protein Global Folds in Perdeuterated Proteins

Ronald A. Venters,† William J. Metzler,‡
Leonard D. Spicer,†,§ Luciano Mueller,‡ and
Bennett T. Farmer II*,‡

Duke University NMR Center and Departments of Biochemistry and Radiology, Duke University Medical Center, P.O. Box 3711, Durham, North Carolina 27710
Bristol-Myers Squibb Pharmaceutical Research Institute P.O. Box 4000, Princeton, New Jersey 08543-4000

Received May 22, 1995

NMR structural studies on isotopically unlabeled proteins are typically limited to those with <100 residues. The advent of $^{13}C/^{15}N$ isotopic enrichment[1−3] and heteronuclear multidimensional experiments[4,5] has greatly extended the size of proteins for which at least sequence-specific resonance assignments can be made.[6,7] A full structural analysis on large proteins still remains problematic. Although 4D heteronuclear NOESY experiments[8−10] facilitate the assignment of NOEs in large proteins,[8] these experiments begin to suffer from a lack of both sensitivity[11] and achievable resolution as the protein size increases.

The perdeuteration of proteins at nonexchangeable sites was originally shown to eliminate many dipolar relaxation pathways.[12,13] Recent work has demonstrated significant increases in both signal-to-noise[14,15] (S/N) and resolution[16] for deuterated proteins in the $^1H_N/^{13}C/^{15}N$ J-correlation experiments.[17,18] Deuteration has also been shown both to enhance the sensitivity[12,13,19] and to simplify the analysis[13,20,21] of NOESY spectra. In this communication, we present an initial analysis of a 4D $^{15}N/^{15}N$-separated NOESY on perdeuterated $^{13}C/^{15}N$-labeled human carbonic anhydrase II[22] (^{2}H-HCA II, 29 kDa monomer) and examine the potential for computing a protein global fold using only $^1H_N-^1H_N$ NOE restraints. The global fold may subsequently be utilized in more detailed structural studies to resolve ambiguous NOE assignments in the 4D $^{13}C/^{13}C$-separated NOESY.

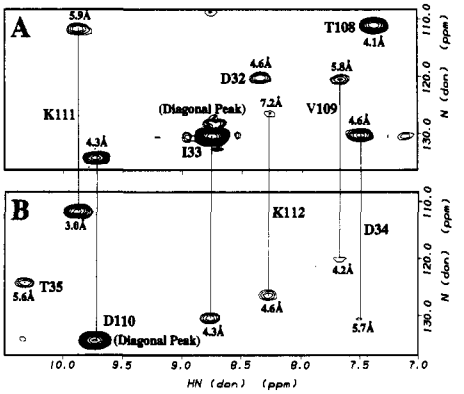

Figure 1. $^1H_{Ndonor}/^{15}N_{donor}$ planes from the 4D $^{15}N/^{15}N$-separated NOESY spectrum on ^{2}H-HCA II. Each peak is labeled with its corresponding distance from the reference structure. (A) The acceptor proton is I33-H_N: (^1H_N,^{15}N) = (8.74,130.2) ppm. The ^{15}N upfield shoulder on the I33-H_N diagonal peak is due to bleed-through from the nearby L239-H_N diagonal peak. (B) The acceptor proton is D110-H_N: (^1H_N,^{15}N) = (9.72,134.0) ppm. The D110−K111 and D110−K112 NOE crosspeaks are intense because K111 begins a turn.

A 7-fold increase in S/N is expected in the 4D $^{15}N/^{15}N$-separated NOESY experiment upon perdeuteration of HCA II. Experimentally, the S/N in a $^1H-^{15}N$ HSQC increases 2.5-fold for ^{2}H-HCA II (unpublished results) and is accounted for by a decrease in the ^1H_N T_2 relaxation rates.[23] Similarly, an additional 70% gain in S/N is calculated to arise from the donor $^1H_N-^{15}N$ INEPT and ^1H_N t_1 evolution periods. A longer ^1H_N T_1 relaxation rate in ^{2}H-HCA II, however, is estimated to decrease the S/N by 15−20% for a 1.27 s recycle time. Finally, the absence of aliphatic protons also minimizes the number of spin-diffusion pathways available to ^1H_N magnetization, thereby allowing the use of longer mixing times.[12,13,19]

A 4D $^{15}N/^{15}N$-separated NOESY was acquired on a 1.6 mM sample of perdeuterated $^{13}C/^{15}N$-labeled HCA II. The pulse sequence and experimental details are provided as supporting information. Although a complete analysis of this 4D spectrum is still in progress, we have identified all sequential $^1H_N-^1H_N$ NOEs in every β-sheet region of ^{2}H-HCA II examined to date. Figure 1 shows a subset of the experimental results for two strands of a β-sheet in ^{2}H-HCA II: NOE cross peaks between the backbone H_N proton of I33 and the backbone H_N protons of intrastrand residues D32 and D34 and interstrand residues T108−K112; and those between the backbone H_N protons of D110 and the backbone H_N protons of intrastrand residues V109, K111, and K112 and interstrand residues I33−T35. In general, direct $^1H_N-^1H_N$ NOEs for distances out to at least 5 Å are readily observed with good S/N.

The process of calculating a protein global fold based solely on $^1H_N-^1H_N$ NOE data has been simulated with two proteins, HCA II and human profilin. For HCA II, a suitable reference structure was generated by the addition of protons onto the crystal structure[24] followed by energy minimization utilizing the program X-PLOR.[25] For human profilin, the refined NMR structure[26] served as the reference structure. Distance restraints were defined by a lower bound of 1.8 Å and an upper bound of $1.25r_{reference}$ (HCA II) and $1.15r_{reference}$ (profilin). A list of $^1H_N-^1H_N$ distance restraints was generated from each reference structure in this manner for the following conditions: (1)

† Duke University NMR Center.
‡ Bristol-Myers Squibb Pharmaceutical Research Institute.
§ Departments of Biochemistry and Radiology, Duke University.

(1) Westler, W. M.; Kainosho, M.; Nagao, H.; Tomonaga, N.; Markley, J. L. *J. Am. Chem. Soc.* **1988**, *110*, 4093−4095.
(2) Ikura, M.; Krinks, M.; Torchia, D. A.; Bax, A. *FEBS Lett.* **1990**, *266*, 155−158.
(3) Venters, R. A.; Calderone, T. L.; Spicer, L. D.; Fierke, C. A. *Biochemistry* **1991**, *30*, 4491−4494.
(4) Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1990**, *89*, 496−514.
(5) Montelione, G. T.; Wagner, G. *J. Magn. Reson.* **1990**, *87*, 183−188.
(6) Fogh, R. H.; Schipper, D.; Boelens, R.; Kaptein, R. *J. Biomol. NMR* **1994**, *4*, 123−128.
(7) Remerowski, M. L.; Domke, T.; Groenewegen, A.; Pepermans, H. A. M.; Hilbers, C. W.; van de Ven, F. J. M. *J. Biomol. NMR* **1994**, *4*, 257−278.
(8) Clore, G. M.; Kay, L. E.; Bax, A.; Gronenborn, A. M. *Biochemistry* **1991**, *30*, 12−18.
(9) Vuister, G. W.; Bax, A. *J. Magn. Reson., Ser. B* **1993**, *101*, 210−213.
(10) Farmer, B. T., II; Mueller, L. *J. Biomol. NMR* **1994**, *4*, 673−687.
(11) Vuister, G. W.; Kim, S.-J.; Wu, C.; Bax, A. *J. Am. Chem. Soc.* **1994**, *116*, 9206−9210.
(12) LeMaster, D. M.; Richards, F. M. *Biochemistry* **1988**, *27*, 142−50.
(13) Torchia, D. A.; Sparks, S. W.; Bax, A. *J. Am. Chem. Soc.* **1988**, *110*, 2320−2321.
(14) Grzesiek, S.; Anglister, J.; Ren, H.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 4369−4370.
(15) Yamazaki, T.; Lee, W.; Arrowsmith, C. H.; Muhandiram, D. R.; Kay, L. E. *J. Am. Chem. Soc.* **1994**, *116*, 11655−11666.
(16) 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.
(17) Grzesiek, S.; Bax, A. *J. Magn. Reson.* **1992**, *99*, 201−207.
(18) Friedrichs, M. S.; Mueller, L.; Wittekind, M. *J. Biomol. NMR* **1994**, *4*, 703−726.
(19) Tsang, P.; Wright, P. E.; Rance, M. *J. Am. Chem. Soc.* **1990**, *112*, 8183−8185.
(20) Arrowsmith, C. H.; Pachter, R.; Altman, R. B.; Iyer, S. B.; Jardetzky, O. *Biochemistry* **1990**, *29*, 6332−6341.
(21) Reisman, J.; Jariel-Encontre, I.; Hsu, V. L.; Parello, J.; Geiduschek, E. P.; Kearns, D. R. *J. Am. Chem. Soc.* **1991**, *113*, 2787−2789.
(22) Venters, R. A.; Huang, C.-C.; Farmer, B. T., II; Trolard, R.; Spicer, L. D.; Fierke, C. A. *J. Biomol. NMR* **1995**, *5*, 339−344.
(23) Markus, M. A.; Dayie, K. T.; Matsudairat, P.; Wagner, G. *J. Magn. Reson., Ser. B* **1994**, *105*, 192−195.

0002-7863/95/1517-9592$09.00/0

backbone H_N protons to only backbone or to both backbone and sidechain (all) H_N protons and (2) $^1H_N-^1H_N$ distance cutoffs of 5, 6, or 7 Å. For HCA II, any side chain restraints known to be unobtainable or ambiguously assigned with current NMR data were removed from the three restraint lists utilizing all H_N protons. For each protein, a family of embedded substructures (all H_N, N, CO, C_α, C_β, C_γ) was calculated for each restraint list using distance geometry.[25] Within each family, structures clearly delineated into two groups based on RMSDs to the reference structure, with one group always having significantly higher RMSDs than the other. The high-RMSD group contained mirror-image structures and was removed from further consideration. Side chains were subsequently built back onto the remaining substructures in each family using a random-coil template; and the resulting structures were regularized using a simulated annealing protocol.[25] These calculated structures were then compared against the relevant reference structure.

For human profilin, restraint lists generated with 5 and 6 Å distance cutoffs using only backbone H_N protons produce structures that are poor with respect to the overall fold. If the cutoff distance is increased to 7 Å, the calculated structures begin to approximate the fold of profilin reasonably well. The C-terminal helix is now aligned properly with the β-sheet. Helix B is well-formed and positioned in the correct general area. The orientation of the N-terminal helix, however, is neither well-formed nor properly aligned with the β-sheet. The restraint list generated with a 5 Å cutoff distance using *all* H_N protons produces structures in which the β-sheet is well-defined; but the relative position of the helices is poorly defined due to a lack of NOE restraints to the β-sheet. When the cutoff distance is increased to 6 Å for this case, the quality of the structures improves dramatically, due in large part to the inclusion of additional NOE contacts to side chain H_N protons. The β-sheet and both N- and C-terminal helices are well-defined; and helix B is positioned in the correct general area. We therefore conclude that the calculation of a reasonable global fold for profilin requires both backbone and side chain $^1H_N-^1H_N$ distance restraints and a minimum cutoff distance of 6 Å.

The same general observations also hold for HCA II. For all non-hydrogen backbone atoms, the RMSDs between the mean structure for each set of conditions and the reference structure are 8.2, 7.2, and 5.7 Å with only backbone ^1H_N restraints for a cutoff distance of 5, 6, and 7 Å, respectively; and similarly 7.9, 6.4, and 4.1 Å with both backbone and side-chain ^1H_N restraints. With the minimum restraint set (5 Å cutoff distance, backbone ^1H_N restraints only), the core β-sheets of HCA II are well-determined and oriented correctly with respect to one another, thereby establishing the majority of the protein global fold. The turns and helices, however, exhibit disorder; and the latter are not properly oriented with respect to the β-sheet core. The N-terminal helix has no restraints orienting it to the rest of the protein and is therefore disordered. Using the restraint list containing both backbone and side chain $^1H_N-^1H_N$ NOEs with a 6 Å cutoff distance tightens up the overall structure; but an ordered placement of helices relative to the β-sheet core continues to be problematic and the N-terminal helix remains disordered. If the cutoff distance is increased to 7 Å, the quality of the structures improves. The N-terminal helix is now both ordered and properly oriented to the rest of the protein; and other helices are correctly positioned with respect to the β-sheet core. Backbone $^1H_N-^1H_N$ NOE restraints alone therefore are insufficient to determine the entire protein global fold for HCA II: restraints involving side chain H_N protons must be identified and assigned. Using all $^1H_N-^1H_N$ NOE restraints, a cutoff distance of 7 Å clearly provides the best determination of the protein global fold; however, even those structures calculated from similar restraints obtained with a 5 Å cutoff distance provide a reasonable fold.

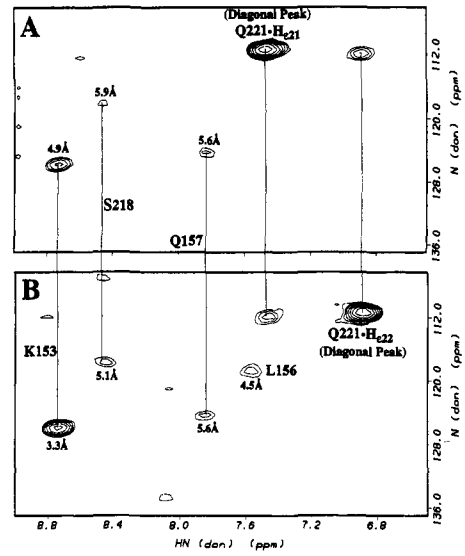

Figure 2. $^1H_{Ndonor}/^{15}N_{donor}$ planes as described in Figure 1. (A) The acceptor proton is Q221-$H_{\epsilon 21}$: (^1H_N,^{15}N) = (6.89,111.5) ppm. (B) The acceptor proton is Q221-$H_{\epsilon 22}$: (^1H_N,^{15}N) = (7.46,111.5) ppm.

The above results indicate that NOE restraints involving both backbone and side chain H_N protons are essential and that a 6 Å cutoff distance is desirable for calculating the protein global fold in perdeuterated proteins. The importance of side chain assignments in obtaining the correct global fold cannot be overemphasized. Part of the improvement in the human profilin structures between the two restraint lists with a 6 Å cutoff distance can be directly attributed to side-chain-to-backbone restraints that serve to orient the N-terminal helix both to the C-terminal helix and to the β-sheet. In ^{2}H-HCA II, experimental NOEs from the side chain H_ϵ protons of Q221, located in an α-helix extending from S218 to L228, to backbone H_N protons of residues in an α-helix extending from K153 to I166 (Figure 2) establish a set of critical distance restraints that serve to define the relative position of these two helices in a global fold calculation. Because the backbone H_N protons of these two helices are separated by over 8 Å on average, the position of these two helices relative to one another cannot be reliably determined from NOE restraints involving backbone H_N protons alone.

We have demonstrated that $^1H_N-^1H_N$ NOE restraint data is sufficient to determine the global fold of a perdeuterated protein provided that (1) restraints to side-chain H_N protons are included and (2) a cutoff distance approaching 6 Å can be attained. We have also shown that NOEs between two H_N protons separated by distances approaching 5.5 Å can be measured utilizing a 1.6 mM perdeuterated $^{13}C/^{15}N$ HCA II sample in a 5 mm probe. A 10-day acquisition on a more concentrated sample (2.5 mM) in an 8 mm probe should extend this cutoff distance to ~7 Å. Additional NOE restraints may be obtainable from sites known to have high fractional protonation in ^{2}H-HCA II, e.g., the C_γ carbon of Glu, Gln, and Arg.[27]

Acknowledgment. Equipment grants from the NIH, NSF, NCBC (Duke NMR Center), and GM41829 (L.D.S.) and comments from Dr. T. N. Jaishree are acknowledged.

Supporting Information Available: Pulse sequence and parameters for 4D $^{15}N/^{15}N$-separated NOESY (3 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions.

JA951640I

(24) Liljas, A.; Kannan, K. K.; Bergsten, P.-C.; Waara, I.; Fridborg, K.; Strandberg, B.; Carlbom, U.; Jarup, L.; Lovgren, S.; Petef, M. *Nature, New Biol.* **1972**, *235*, 131–137.

(25) Brunger, A. T. *X-PLOR version 3.1*; Yale University Press: New Haven, CT, 1992.

(26) Metzler, W. J.; Farmer, B. T., II; Constantine, K. L.; Friedrichs, M. S.; Lavoie, T.; Mueller, L. *Protein Sci.* **1995**, *4*, 450–459.

(27) Farmer, B. T., II; Venters, R. A. *J. Biomol. NMR*, submitted for publication.