J|A|C|S
COMMUNICATIONS

Published on Web 03/07/2006

Protonless NMR Experiments for Sequence-Specific Assignment of Backbone Nuclei in Unfolded Proteins

Wolfgang Bermel,[†] Ivano Bertini,*[,‡] Isabella C. Felli,[‡] Yong-Min Lee,[‡] Claudio Luchinat,[§] and Roberta Pierattelli[‡]

Bruker BioSpin GmbH, Rheinstetten, Germany, CERM and Department of Chemistry, University of Florence, Italy, and CERM and Department of Agricultural Biotechnology, University of Florence, Italy

Received December 3, 2005; E-mail: ivanobertini@cerm.unifi.it

There is increasing evidence that a non-negligible number of proteins may be natively unfolded, or that unfolded or partially unfolded states for otherwise folded proteins may be relevant in particular phases of cell cycle.[1,2] Therefore, characterization of these proteins at the molecular level—and subsequently of their interactions with other cellular or extracellular components—is becoming an increasingly important issue.

At variance with well-folded proteins, crystallization of unfolded proteins is almost impossible, so NMR is the only technique to obtain detailed structural and dynamic information. In principle, NMR is equally well-suited for folded and unfolded proteins, the latter often showing even sharper signals than the former. However, chemical shift dispersions in unfolded proteins are much smaller, causing problems in resonance assignment.[3−6]

Indeed, for backbone nuclei, the chemical shift is roughly a sum of an environment-independent contribution that is characteristic for each amino acid and sensitive to its neighbors in the sequence, and an environment-dependent contribution that is sensitive to the secondary and tertiary structure elements that surround in space that particular nucleus/residue. The latter contribution is virtually absent in an unfolded protein. The environment-independent contribution for H^N and H^α protons is scarcely sensitive to the type of amino acid to which the proton belongs, whereas it is still strongly sensitive for heteronuclei, in the order $C^\alpha < C' < N$.[7−9] Thus, the chemical shift dispersion in the proton dimension is lost, and recording purely heteronuclear spectra becomes a strategy to minimize dimensionality and maximize dispersion. A complete set of experiments based on *protonless* ^{13}C direct detection has been recently proposed for heteronuclear assignment[10−13] with the aim of allowing the investigation of large[14−16] or paramagnetic proteins.[17−21] We here develop a protonless strategy for the study of unfolded proteins. It relies on combining the 3D CBCACON−IPAP experiment[13] for the identification of spin systems with a novel 3D COCON−IPAP experiment (Figure SI1 in the Supporting Information) to obtain the sequence-specific assignment. With respect to analogous ^1H-based experiments,[5,22−25] the favorable chemical shift dispersion of carbonyls[5,6] is exploited in the direct acquisition dimension (see Figure SI2 for a comparison of signal dispersion for the two nuclei). The case study is the natively unfolded, 140 amino acid protein α-synuclein.[26−28] Its backbone assignment was obtained[28−30] by using extended sets of proton-based experiments, such as HNCA, HN(CO)CA, HNCO, HN(CA)CO, CBCANH, and CBCA(CO)NH.[28] The present *two* experiments represent an elegant strategy that provides aesthetically excellent spectra and leads to a sound assignment overcoming any overlap problem, despite the intrinsic lower sensitivity of ^{13}C with respect to ^1H nuclei. The virtually complete assignment for ^{15}N, $^{13}C'$, $^{13}C^\alpha$, and $^{13}C^\beta$ is reported as Supporting Information (Table SI1).

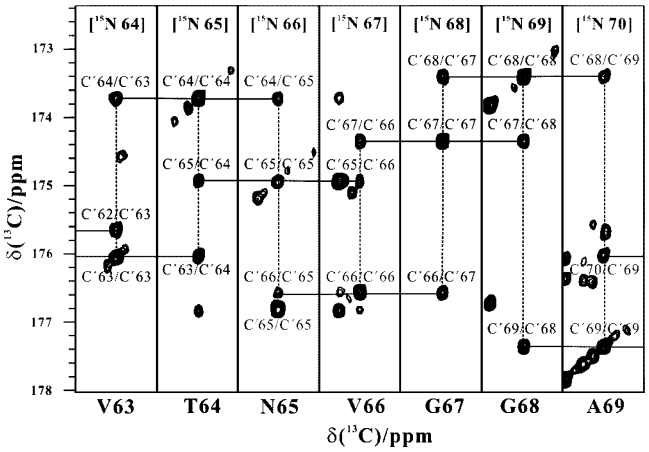

Figure 1. Representative slices of $^{13}C'-^{13}C'$ planes at specific ^{15}N resonances taken from the 16.4 T COCON−IPAP spectrum of $^{13}C-^{15}N$-labeled α-synuclein at 288 K. For each residue indicated at the bottom, the three correlations necessary for sequence-specific assignment ($C'_i-C'_i-N_{i+1}$, $C'_{i-1}-C'_i-N_{i+1}$, $C'_{i+1}-C'_i-N_{i+1}$) are shown (as well as the correlations with the side chain of Asn 65).

The 3D COCON−IPAP correlates the backbone nitrogen with the attached carbonyl carbon and with the previous and following carbonyl carbons in the sequence. The relevant correlations for sequence-specific assignment can then be identified in the $^{13}C'-^{13}C'$ planes at specific ^{15}N resonances. The use of the IPAP method to remove the C′ signal splitting due to the presence of $C'-C^\alpha$ coupling[11] permits also to obtain optimal resolution.

As shown in Figure 1, for each carbonyl carbon, it is indeed possible to identify three peaks: the "diagonal" peak ($C'_i-C'_i-N_{i+1}$), the correlation with the previous ($C'_{i-1}-C'_i-N_{i+1}$), and the correlation with the following ($C'_{i+1}-C'_i-N_{i+1}$) backbone carbonyl carbons. The transfer of magnetization from a carbonyl carbon to both the previous and the following provides redundant information, always welcome for assignment purposes. The ambiguity in the direction of the sequential assignment is easily removed by using any of the other protonless experiments.[31] In this strategy, we use the 3D CBCACON−IPAP experiment[13] (Figure 2) that allows us to link each backbone nitrogen with the attached C′ and with the intraresidue C^α (and C^β). In most cases, the amino acid type is identified from the information retained in the chemical shift values themselves.[32] The complete $^{13}C'$ and ^{15}N backbone assignment is reported in Figure 3 by labeling the correlations detected in the 2D CON−IPAP experiment,[13] as well as in Table SI1 together with the $^{13}C^\alpha$ and $^{13}C^\beta$ assignment. Figure 3 also demonstrates the excellent spreading achievable with a 2D experiment for an unfolded protein, provided the proton dimension is left out.

Signal intensities in the COCON−IPAP experiment can be related to the $^3J_{C'C'}$ value, providing a source of structural information. In the present case, we noticed that one of the three connectivities of the proline C′ carbons is weaker than that of the

[†] Bruker BioSpin GmbH.
[‡] CERM and Department of Chemistry, University of Florence.
[§] CERM and Department of Agricultural Biotechnology, University of Florence.

10.1021/ja0582206 CCC: $33.50

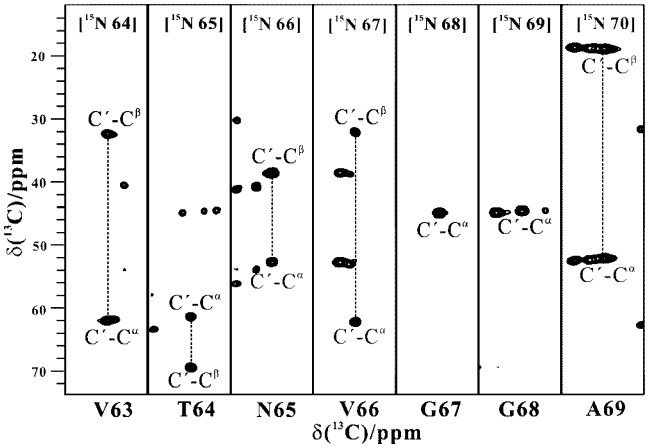

Figure 2. Representative slices of $^{13}C^{ali}-^{13}C'$ planes at specific ^{15}N resonances taken from the 16.4 T CBCACON−IPAP spectrum of $^{13}C-^{15}N$-labeled α-synuclein at 288 K. For each residue, the $C^{\alpha}_i-C'_i-N_{i+1}$ and, for non-Gly residues, the $C^{\beta}_i-C'_i-N_{i+1}$ correlations are shown.

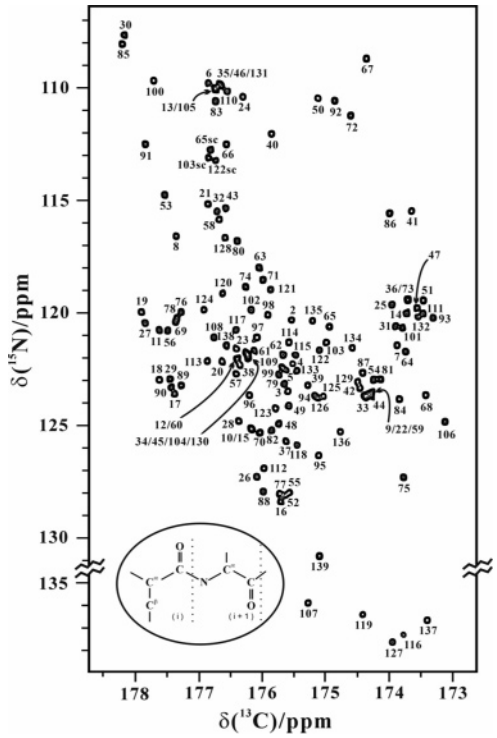

Figure 3. The 16.4 T CON−IPAP spectrum of $^{13}C-^{15}N$-labeled α-synuclein at 288 K. The signal assignment obtained through the COCON−IPAP experiment is reported next to each C'_i-N_{i+1} cross-peak.

others, and the same holds for the $(i+1)$ connectivity of the residue preceding a Pro (see Figure SI3). This is due to the typical values the ϕ dihedral angle can assume in prolines[33] that results in a fairly small $^3J_{C'C'}$ value. The connectivities of all the other amino acids show similar intensities along the sequence, indicating extensive conformational averaging and thus the absence of a pronounced secondary structure propensity of the protein in solution.[5,6] The resulting averaged $^3J_{C'C'}$ values, despite small, are still able to provide measurable peaks due to the robustness of isotropic mixing.[34,35] Furthermore, the averaged $^3J_{C'C'}$ values are fairly constant,[5,6,24] so that the COCON−IPAP experiment is the experiment of choice for unfolded proteins because the observability of all cross-peaks (including those of prolines) should be guaranteed.

Acknowledgment. This work has been supported in part by the EC Contracts QLG2-CT-2002-00988 and 005032. The NMR pulse sequences written for the Bruker Avance series spectrometers are available upon request. The signals assignment is deposited in BMRB (entry 6968).

Supporting Information Available: Description of the 3D COCON−IPAP sequence; experimental NMR parameters used; comparison of signal dispersion for $^1H^N$ and $^{13}C'$; slices of selected planes of the 3D COCON−IPAP experiment containing Pro signals; ^{15}N, $^{13}C'$, $^{13}C^{\alpha}$, and $^{13}C^{\beta}$ assignments of $^{13}C-^{15}N$ α-synuclein. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Dobson, C. M. *Nature* **2003**, *426*, 884−890.
(2) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197−208.
(3) Zhang, O.; Forman-Kay, J. D.; Shortle, D.; Kay, L. E. *J. Biomol. NMR* **1997**, *9*, 181−200.
(4) Hennig, M.; Bermel, W.; Spencer, A.; Dobson, C. M.; Smith, L. J.; Schwalbe, H. *J. Mol. Biol.* **1999**, *288*, 705−723.
(5) Liu, A.; Riek, R.; Wider, G.; Von Schroetter, C.; Zahn, R.; Wüthrich, K. *J. Biomol. NMR* **2000**, *16*, 127−138.
(6) Dyson, H. J.; Wright, P. E. *Methods Enzymol.* **2001**, *339*, 258−270.
(7) Braun, D.; Wider, G.; Wüthrich, K. *J. Am. Chem. Soc.* **1994**, *116*, 8466−8469.
(8) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67−81.
(9) (a) Yao, J.; Dyson, H. J.; Wright, P. E. *FEBS Lett.* **1997**, *419*, 285−289. (b) Schwarzinger, S.; Kroon, G. J. A.; Foss, T. R.; Chung, J.; Wright, P. E.; Dyson, H. J. *J. Am. Chem. Soc.* **2001**, *123*, 2970−2978.
(10) Bertini, I.; Duma, L.; Felli, I. C.; Fey, M.; Luchinat, C.; Pierattelli, R.; Vasos, P. R. *Angew. Chem., Int. Ed.* **2004**, *43*, 2257−2259.
(11) Bermel, W.; Bertini, I.; Duma, L.; Emsley, L.; Felli, I. C.; Pierattelli, R.; Vasos, P. R. *Angew. Chem., Int. Ed.* **2005**, *44*, 3089−3092.
(12) Bermel, W.; Bertini, I.; Felli, I. C.; Pierattelli, R.; Vasos, P. R. *J. Magn. Reson.* **2005**, *172*, 324−328.
(13) Bermel, W.; Bertini, I.; Felli, I. C.; Kümmerle, R.; Pierattelli, R. *J. Magn. Reson.* **2005**, *177*, 404−412.
(14) Bertini, I.; Felli, I. C.; Kümmerle, R.; Moskau, D.; Pierattelli, R. *J. Am. Chem. Soc.* **2004**, *126*, 464−465.
(15) Bertini, I.; Felli, I. C.; Kümmerle, R.; Luchinat, C.; Pierattelli, R. *J. Biomol. NMR* **2004**, *30*, 245−251.
(16) Pervushin, K.; Eletsky, A. *J. Biomol. NMR* **2003**, *25*, 147−152.
(17) Bertini, I.; Luchinat, C.; Parigi, G.; Pierattelli, R. *ChemBioChem* **2005**, *6*, 1536−1549.
(18) Bermel, W.; Bertini, I.; Felli, I. C.; Kümmerle, R.; Pierattelli, R. *J. Am. Chem. Soc.* **2003**, *125*, 16423−16429.
(19) Arnesano, F.; Banci, L.; Bertini, I.; Felli, I. C.; Luchinat, C.; Thompsett, A. R. *J. Am. Chem. Soc.* **2003**, *125*, 7200−7208.
(20) Machonkin, T. E.; Westler, W. M.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 3204−3205.
(21) Kostic, M.; Pochapsky, S. S.; Pochapsky, T. C. *J. Am. Chem. Soc.* **2002**, *124*, 9054−9055.
(22) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 185−204.
(23) Muhandiram, D. R.; Kay, L. E. *J. Magn. Reson. Ser. B* **1994**, *103*, 203−216.
(24) Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 8170−8171.
(25) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1997**, *9*, 207−211.
(26) Apo $^{13}C-^{15}N$ α-synuclein was purchased from Protera srl (Sesto Fiorentino, Italy) and used without further purification. The NMR sample contained 1.0 mM α-synuclein, 200 mM KCl, 10 mM 2-morpholinoethanesulfonic acid in 90% H_2O/10% D_2O at pH 6.8.
(27) Weinreb, P. H.; Zhen, W. G.; Poon, A. W.; Conway, K. A.; Lansbury, P. T., Jr. *Biochemistry* **1996**, *35*, 13709−13715.
(28) Eliezer, D.; Kutluay, E.; Bussell, R., Jr.; Browne, G. *J. Mol. Biol.* **2001**, *307*, 1061−1073.
(29) Fernandez, C. O.; Hoyer, W.; Zweckstetter, M.; Jares-Erijman, E. A.; Subramaniam, V.; Griesinger, C.; Jovin, T. M. *EMBO J.* **2004**, *23*, 2039−2046.
(30) Dedmon, M. M.; Lindorff-Larsen, K.; Christodoulou, J.; Vendruscolo, M.; Dobson, C. M. *J. Am. Chem. Soc.* **2005**, *127*, 476−477.
(31) Bermel, W.; Bertini, I.; Felli, I. C.; Piccioli, M.; Pierattelli, R. *Prog. NMR Spectrosc.* **2005**, in press.
(32) Zhang, H.; Neal, S.; Wishart, D. S. *J. Biomol. NMR* **2003**, *25*, 173−195.
(33) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95−99.
(34) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521−528.
(35) As discussed also in ref 5, the $^3J_{C'C'}$ values in an unfolded protein are of the order of the hertz. Accordingly, the mixing time to be used in a TOCSY experiment is rather long (190 ms in the present case). However, the losses due to relaxation during the timing of the sequence in unfolded proteins are less severe that those in folded proteins, and most of the signal intensity is retained.

JA0582206