Protein-Structure Determination

A Heteronuclear Direct-Detection NMR Spectroscopy Experiment for Protein-Backbone Assignment**

Ivano Bertini, Luminita Duma, Isabella C. Felli, Michael Fey, Claudio Luchinat, Roberta Pierattelli, and Paul R. Vasos*

The starting point of any standard protocol for protein structure determination by NMR spectroscopy is the sequence-specific resonance assignment of the polypeptide chain. The established procedure for obtaining sequential assignments involves uniform $^{13}C/^{15}N$ labeling and the exploitation of heteronuclear scalar couplings with ^{1}H-detected triple-resonance experimental schemes. However, ^{1}H nuclei detection in large macromolecules suffers from considerable line broadening. This adverse effect can be alleviated with the use of TROSY in either partially or uniformly deuterated proteins[1–3] in combination with cross relaxation-induced polarization transfer (CRIPT) and cross relaxation-enhanced polarization transfer (CRINEPT) techniques.[4,5]

A valuable alternative to overcome fast relaxation is the use of ^{13}C-nuclei direct detection, which takes advantage of the slowly relaxing ^{13}C magnetization compared to ^{1}H nuclei. Most of the NMR spectroscopic experiments for heteronuclear backbone assignment can, in principle, be performed by using direct heteronuclear excitation and detection without involving protons at any place in the sequence. However, the design strategy of the experiments for backbone assignment by using heteronuclei only is different from that for ^{1}H-based experiments. Indeed, most of the experiments used nowadays for backbone assignment are based on the HN group, which actively exploits TROSY effects.[6,7] Passing to direct heteronuclear detection, ^{15}N nuclei are not the best starting point for sensitivity reasons. The experiments should rather start and end on ^{13}C nuclei, either on C^{α} or on carbonyls C' atoms, in an out-and-back scheme,[8] or should start on one of them and end on the other in an out-and-stay approach.[9] Among the possible schemes, the latter gave the best results and its implementation as a 3D experiment, which we named CANCO, to accomplish the sequence-specific assignment of a protein is here presented.

The 3D CANCO correlates the chemical shifts of C^{α} nuclei with the shifts of the two neighboring nitrogen nuclei and subsequently with those of the carbonyls adjacent to these nitrogen atoms. This produces patterns in which for each N spin two resonances appear, which correspond to the two nearest C^{α} carbon atoms (C^{α}_i, C'_i, N_{i+1} and C'_i, N_{i+1},C^{α}_{i+1}). In the 2D COCA spectrum only the intraresidue correlation (C^{α}_i, C'_i) is present, and therefore backbone assignment can be performed with the above pair of experiments. The pulse sequence follows the transfer scheme reported in Figure 1.

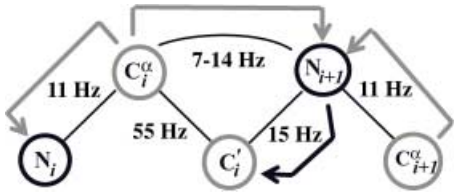

Figure 1. Transfer pathway in the CANCO sequence: the transfers that occur during the first step of the sequence are indicated by gray arrows, those effected during the second step by black arrows.

The description of the actual pulse sequence and the magnetization transfer pathway in terms of product operator formalism are detailed in the Supporting Information.

The first selective pulse on C^{α} is followed by a constant-time (CT) evolution period, optimized to refocus $C^{\alpha}C^{\beta}$ couplings, during which magnetization is labeled with the C^{α} chemical shift. Simultaneously, as the coupling to CO is refocused, transfer to the two nearest nitrogen spins is achieved by allowing the two $C^{\alpha}N$ couplings to evolve during part of the CT period. In the second CT evolution period, the chemical-shift labeling of the nitrogen spins is accomplished, the magnetization is refocused with respect to C^{α} and *J*-coupling transfer to carbonyls takes place. The last part of the sequence consists of refocusing of CO magnetization with respect to nitrogen atoms and detection on carbonyls. The concatenation of the N-—CO transfer with the refocusing of ^{15}N magnetization with respect to C^{α}, allowed by the "out-and-stay" transfer approach chosen, considerably reduces the duration of the sequence, thus ensuring an increase in sensitivity.[*]

[*] Prof. I. Bertini, Dr. I. C. Felli, Prof. R. Pierattelli
Magnetic Resonance Center and
Department of Chemistry
Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence (Italy)
Fax: (+39) 055–457–4271
E-mail: bertini@cerm.unifi.it
L. Duma
Laboratoire de Chimie
Ecole Normale Superieure de Lyon (France)
Dr. M. Fey
Bruker BioSpin AG
Fällanden (Switzerland)
Prof. C. Luchinat
Magnetic Resonance Center and
Department of Agricultural Biotechnology
University of Florence (Italy)
P. R. Vasos
Magnetic Resonance Center
University of Florence (Italy)

[**] L.D. was supported by the EC Marie Curie Fellowship, Contract Number HPMT-2000–000137. P.R.V. acknowledges the "CROSS-CORRELATION" Research Training Network, Contract Number HPRN-CT-2000–00092. The authors thank Dr. Wolfgang Bermel for critical reading of the manuscript and Prof. Lyndon Emsley for stimulating discussions and useful comments.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

[*] The starting C^{α} magnetization can be enhanced by irradiating protons during part of the recycle delay, which results in an increment in sensitivity that depends on the C—H cross relaxation rate. This method, however, will not be useful in ^{2}H labeled molecules.

 DOI: 10.1002/ange.200453661

The initial CT delay was set to $1/J_{C^\alpha C^\beta}$ to refocus the $C^\alpha C^\beta$ coupling. An alternative is to use a selective 180° pulse at the C^α frequency (with respect to the C^β frequency), which results in a decreased sensitivity of amino acids for which the C^α and C^β shifts are very similar (Thr, Ser) or of amino acids for which the C^α resonances have large deviations from the mean C^α chemical shift of proteins (Gly).[10] Phase cycling eliminates artefacts due to imperfections of the associated 180° pulse and coherences resulting from incomplete *J*-coupling evolution. The pulsed field gradients are complementary to the phase cycling, thus eliminating any in-plane magnetization once the intended *J*-coupling transfers are achieved and the magnetization of interest is parallel to the main field.

The 3D CANCO experiment was applied to the analysis of oncomodulin as a test case. Oncomodulin is a small Ca-binding protein (109 amino acids), which has been recently investigated in our laboratory and for which a virtually complete assignment is available (100% C^α, 99.1% CO and 98.2% backbone ^{15}N).[11] A pair of 2D experiments, namely the COCA[12] and CON[13] experiments, were also acquired to complement the 3D CANCO data (see Supporting Information).

Figure 2 illustrates the sequential assignment approach for the three residues 21Asp–22Pro–23Asp. The ^{15}N chemical shifts of the two prolines present in oncomodulin (22P, at 136.5 ppm and 27P, at 135.6 ppm) were actually assigned by using this 3D spectrum. Indeed, all types of residues can be detected with the 3D CANCO sequence, in contrast with other sequences that, being based on the NH group, do not allow a direct assignment of proline residues. The 3D experiment displays the signals of 99 C'_i, N_{i+1},C^α_{i+1} correlations (92.5% of the assigned expected correlations) and 85 C^α_i, C'_i, N_{i+1} correlations (79.4% of the assigned expected correlations). All missing C^α_i, C'_i correlations were identified in the 2D COCA experiment and were connected to the adjacent residue with the aid of a 2D CON experiment.

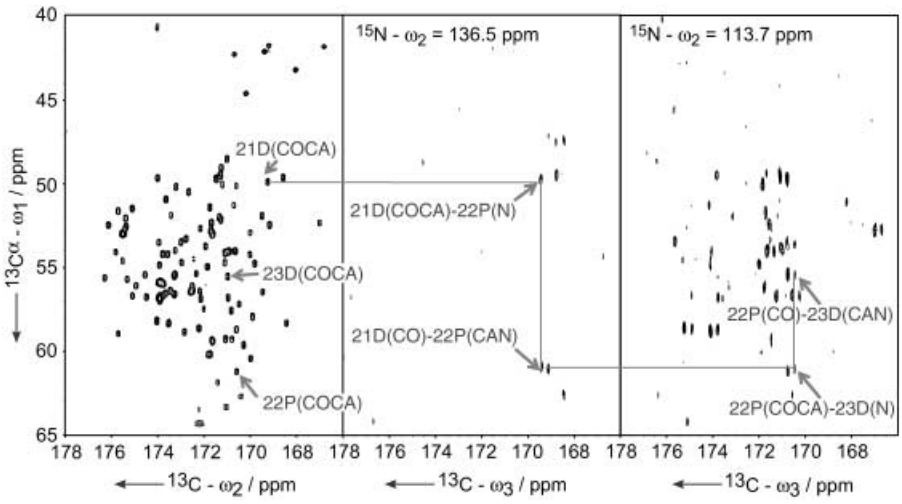

Figure 2. Schematic representation of the methodology for establishing sequential correlations in the series of residues 21 Asp–22 Pro–23 Asp of oncomodulin, by using the 2D CACO and the 3D CANCO spectrum. Two planes of the 3D spectrum are displayed, corresponding to the ^{15}N frequency of 22 Pro and 23 Asp.

The ^{15}N resonances of 22 P (136.5 ppm), 27 P (135.6 ppm), 52 D (118.6 ppm), 64 L (116.6 ppm), 65 K (116.9 ppm) and 109 S (122.0 ppm), the $^{13}C^\alpha$ resonances of 59 L (58.8 ppm), 60 D (56 ppm), 63 E (55.48 ppm), 64 L (54 ppm) and 109 S (58.09 ppm), as well as the CO resonances of 64 L (175.2 ppm) and 86 L (174.4 ppm) were not assigned by the use of standard proton spectra [15] and could be assigned by using the ^{13}C-detected spectra described here.

A characteristic of the sequence is its increased sensitivity for interresidue correlations, which are favored with respect to intraresidue correlations, as the former originate from transfer with the one bond $C^\alpha N$ coupling constant (generally higher), while the latter originate from magnetization transfer with the two bonds $C^\alpha N$ coupling constant (generally smaller). This is a remarkable feature, as in the majority of the existing proton pulse sequences that provide both correlations simultaneously the interresidue correlations are unfavorable, thus making it difficult to establish sequential connectivities when the quality of the spectra is poor.

Resonances that involve a step in magnetization transfer on the proline nitrogen spins (intra- and intercorrelations for residues preceding prolines) are favored in terms of sensitivity, as the relaxation rates of proline nitrogen nuclei are lower (they are coupled to an additional carbon atom, while all other nitrogen atoms are coupled to a proton). On the other hand, the sensitivity is lower when transfer from C^α of glycines is involved, as these nuclei are coupled to two protons, but this is a concern also for the classical proton-based sequences. In the 3D spectrum recorded on the oncomodulin sample, out of the 12 overall expected correlations originating on Gly C^α nuclei, only three C^α_i, C'_i, N_{i+1} and two C'_i, N_{i+1},C^α_{i+1} correlations were observed. On the other hand, due to the absence of the coupling with C^β during the first CT evolution period, signals originating from Gly C^α are very easily distinguished, as they have opposite phase with respect to all other signals that evolve with the cosine period during this step.

Magnetization that originates on Ala C^α nuclei might also result in less sensitive transfers, because the chemical shifts of C^α and C^β of alanine are at the extreme ranges of the irradiated region.[10] The remaining three missing intercorrelations might be accounted for by slightly different values of the interresidual coupling between nitrogen and C^α carbon atoms (see Supporting Information), for which the transfer delay was not optimal, or by the occurrence of conformational exchange phenomena. The magnetization-transfer pathway hinders the observation of an intraresidue correlation for the last residue in a protein (as magnetization would have to be transferred by the nitrogen nucleus of the next residue), but the interresidue correlation between the penultimate and the last residue is

expected, as well as both the correlations involving the first residue.

The triple-resonance experiment presented here displays information about the chemical shift for all three backbone heteronuclei of an amino acid, thus enabling the complete assignment of the chemical shift of heteronuclei in a protein by use of only two experiments such as the 3D CANCO and a 2D COCA. This is in contrast with the classical proton-based approach, in which four 2D/3D (e.g. HNCO, HN(CA)CO, HNCA, HN(CO)CA) experiments are necessary to obtain the assignment of all the backbone heteronuclei in a polypeptide chain.[14] Clearly, in the latter case there is the advantage to have also the assignment of all the peptidic protons. However, this advantage is lost when line broadening prevents coherence transfer to heteronuclei.

The sensitivity of the 3D CANCO spectrum can be compared with that of the HN(CA)CO spectrum, provided we take into account relaxation effects during transfer and detection periods, as well as the field and gamma dependency of sensitivity. For a protein that has a 10 ns rotational correlation time, such as the oncomodulin sample at 283 K used here, simulations indicate that the proton sequence is 2.5 times more sensitive than the carbon-detected sequence. However, as the rotational correlation time increases (e.g. 20 ns) transfer periods in which proton magnetisation is in-plane considerably reduce the relative sensitivity of the proton-detected sequence (e.g. about 10%).

In the absence of TROSY effects, which become operative at high fields, the present approach becomes even more appealing. In paramagnetic molecules, the TROSY effect may be ineffective in some parts of the macromolecule, due to paramagnetic relaxation, thus the heteronuclear direct detection sequences are preferable.

Received: January 2, 2004 [Z53661]
Published Online: March 22, 2004

Keywords: NMR spectroscopy · protein structures · proteins · structure elucidation

[1] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12366 – 12371.
[2] M. Salzmann, K. Pervushin, G. Wider, H. Senn, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13585 – 13590.
[3] M. Salzmann, G. Wider, K. Pervushin, H. Senn, K. Wüthrich, *J. Am. Chem. Soc.* **1999**, *121*, 844 – 848.
[4] R. Riek, G. Wider, K. Pervushin, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4918 – 4923.
[5] R. Riek, J. Fiaux, E. B. Bertelsen, A. L. Horwich, K. Wüthrich, *J. Am. Chem. Soc.* **2002**, *124*, 12144 – 12153.
[6] K. Pervushin, *Q. Rev. Biophys.* **2000**, *33*, 161 – 197.
[7] C. Fernandez, G. Wider, *Curr. Opin. Struct. Biol.* **2003**, *13*, 570 – 580.
[8] S. J. Archer, M. Ikura, D. A. Torchia, A. Bax, *J. Magn. Reson.* **1991**, *95*, 636 – 641.
[9] L. E. Kay, M. Ikura, A. Bax, *J. Magn. Reson.* **1991**, *91*, 84 – 92.
[10] D. S. Wishart, B. D. Sykes, *J. Biomol. NMR* **1994**, *4*, 171 – 180.
[11] ^{13}C, ^{15}N labeled oncomodulin expression and purification will be presented elsewhere. The protein concentration in the final NMR samples was about 2.5 mM. The sample was dissolved in water solution of 100 mM NaCl at pH 6.0 with a 10% D_2O added for the field lock. ^{13}C nuclei direct detection experiments were carried out on a 16.4 T Bruker AVANCE 700 spectrometer, operating at 700.06 MHz for ^{1}H and 176.03 MHz for ^{13}C and equipped with a prototype TXO ^{13}C, {^{15}N, ^{1}H} probe. All spectra were recorded at 283 K.
[12] W. Bermel, I. Bertini, I. C. Felli, R. Pierattelli, R. Kümmerle, *J. Am. Chem. Soc.* **2003**, *125*, 16423 – 16429.
[13] F. Arnesano, L. Banci, I. Bertini, I. C. Felli, C. Luchinat, A. R. Thompsett, *J. Am. Chem. Soc.* **2003**, *125*, 7200 – 7208.
[14] M. Ikura, L. E. Kay, A. Bax, *Biochemistry* **1990**, *29*, 4659 – 4667.
[15] I. Bertini, Unpublished results.

