J|A|C|S COMMUNICATIONS

Published on Web 08/24/2009

A Double TROSY hNCAnH Experiment for Efficient Assignment of Large and Challenging Proteins

Dominique P. Frueh,*,† Haribabu Arthanari,† Alexander Koglin,†,‡ Christopher T. Walsh,† and Gerhard Wagner*,†

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, and Los Alamos National Laboratory, Bioenergy & Environmental Science, Bioscience Division, B-8, Los Alamos, New Mexico 87545

Received June 8, 2009; E-mail: Gerhard_Wagner@hms.harvard.edu; dominique_frueh@hms.harvard.edu

Nuclear magnetic resonance (NMR) is a versatile technique providing structural, dynamic, and kinetic information on a variety of biological systems. All these applications rely on the ability to assign the signals observed in the NMR spectra to the nuclei of the molecule under investigation. For backbone nuclei of proteins, this is traditionally achieved by recording a series of triple resonance experiments that each correlate an amide proton−nitrogen pair with carbon nuclei belonging to sequentially adjacent residues (e.g., $C^{\alpha}_{(i)}$ and $C^{\alpha}_{(i-1)}$).[1] Strips along the carbon dimension are generated and are then compared for matching carbon frequencies. In many challenging systems, such as α-helical or unfolded proteins or molecules with undiversified amino acid sequences (e.g., ...QQEQQ... or SDLA...SDLA...), this task becomes tedious and often unsuccessful due to spectral crowding and resonance degeneracies, in particular for carbon nuclei. In this case, other methods are used to supplement or supplant the conventional suite of experiments. Among these are the hNcocaNH and HncocaNH pair of experiments (aka (H)N(COCA)NNH[2] or HN(C)N[3] and H(NCOCA)NNH[4] or HNN[3]), which allow for a simple and efficient "H^N−N backbone walk" assignment procedure that directly correlates H^N and N nuclei.[5] However, this necessitates the acquisition of two NMR spectra. A recently described alternative method exploits redundancies in the HNCA and HACAN experiments to provide a "stairway" assignment procedure in which chain extension of sequential residues is accomplished by navigating between the $^1H-^{13}C$ planes of the 3D spectrum.[6] This hybrid spectrum can be obtained with a single time-shared experiment. Resonance degeneracy is also a limitation for large proteins, which feature an increased number of residues. Large proteins in addition suffer from faster relaxation resulting in line broadening and signal losses during spin manipulations and evolution periods in the experiments. Thus, while additional experiments such as those described above are needed to overcome ambiguities in assignment, their application is limited by these constraints.

Here, we present a novel NMR experiment that exhibits visually prominent spectral patterns and enables facile sequential backbone assignments in a single experiment. The hNCAnH pulse sequence is a modification of the recently developed TROSY-hNcaNH-TROSY,[7] which has been used successfully to assign large proteins.[8] The major difference between these two experiments is that the chemical shift evolution of nitrogens attached to the detected nuclei (which would provide an HN-HSQC type of correlation) is replaced by a C^{α} constant time evolution. This modification creates easily recognizable rectangular spectral patterns in the $^1H-^{15}N$ planes and allows for easy and unambiguous sequential assignment of H^N, N, and C^{α} nuclei in a single experiment. Several versions of the pulse sequence, appropriate for proteins of different sizes, are discussed in the Supporting Information. For large proteins, the new experiment was designed to overcome relaxation losses by using the double TROSY method,[7] which minimizes transverse relaxation not only for a pair of chemically bound H^N and N nuclei, as in conventional TROSY, but also for nuclei that are linked by a 3D cross peak but are not directly connected by a single scalar coupling. This results in sensitivity which is approximately half that of the HNCACB experiment. The magnetization flow can be summarized as follows (see Supporting Information):

$$\left.\begin{array}{l} H_{i-1} \rightarrow N_{i-1}(t_1) \xrightarrow{T_{N'}} C_{i-1}(t_2) \xrightarrow{T_C} N_i \xrightarrow{T_N} \\ H_i \rightarrow N_i(t_1) \left\{\begin{array}{l} \xrightarrow{T_{N'}} C_{i-1}(t_2) \xrightarrow{T_C} N_i \xrightarrow{T_N} \\ \xrightarrow{T_{N'}} C_i(t_2) \xrightarrow{T_C} N_i \xrightarrow{T_N} \end{array}\right. \\ H_{i+1} \rightarrow N_{i+1}(t_1) \xrightarrow{T_{N'}} C_i(t_2) \xrightarrow{T_C} N_i \xrightarrow{T_N} \end{array}\right\} \rightarrow H_i(t_3)$$

Here, H denotes H^N and C denotes C^{α}. A summary of the coherence flow between correlated nuclei is shown in Figure 1.

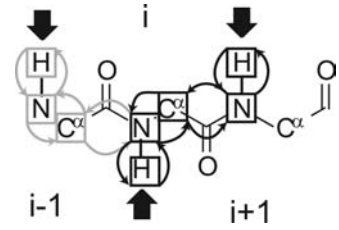

Figure 1. Schematic representation of the transfers involved in the hNCAnH experiments. Only pathways leading to detection of the amide proton of the i^{th} residue are considered. The squares indicate the coherences that are frequency labeled in the experiments. Black arrows and squares denote transfers and evolution periods that relate residues i and i+1. Gray arrows denote transfers that relate residues i and i−1.

Offsetting all source protons by one residue (Supporting Information) reveals that, at a given carbon frequency, four cross-peaks arranged in a rectangular pattern are to be expected in each H/N plane: (H^N_i,N_i,C^{α}_i), (H^N_i,N_{i+1},C^{α}_i), (H^N_{i+1},N_{i+1},C^{α}_i), and (H^N_{i+1},N_i,C^{α}_i). Sequential correlation peaks along both H and N dimensions provide unambiguous identification of proton and nitrogen signals of successive residues in a manner reminiscent of NOESY cross-peaks (H^N−N backbone walk). Since sequential information also appears in the carbon dimension, a cross section along C at (H^N_i,N_i) enables the identification of the C^{α} frequency of the previous residue (stairway). The H/N plane at the resulting C^{α}_{i-1} frequency can then be analyzed as described above, leading to the assignment of H^N_{i-1} and N_{i-1}.

The advantages of the method are illustrated with the 33 kDa, 297 residue component B of the *E. coli* enterobactin synthetase.[9]

† Harvard Medical School.
‡ Los Alamos National Laboratory.

10.1021/ja9046685 CCC: $40.75

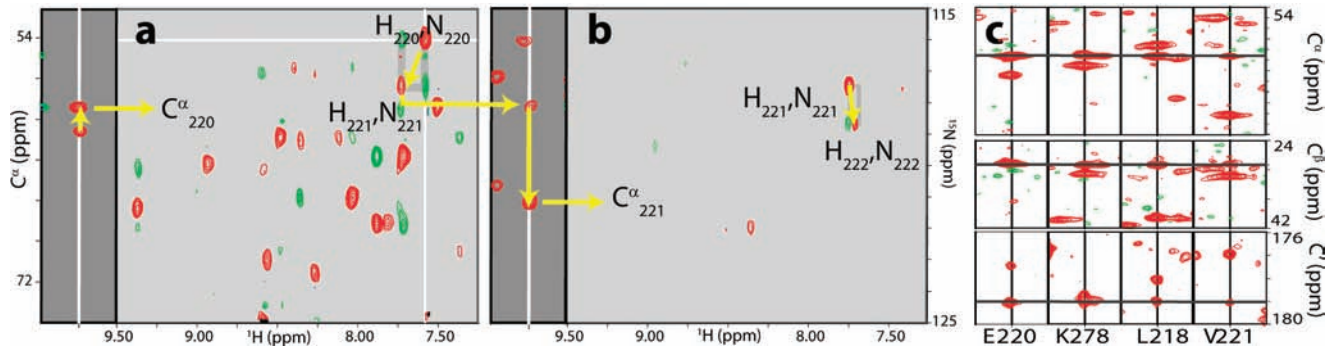

Figure 2. (a, b) Stairway backbone walk. Intraresidue (H, N) correlations are positive (red) while inter-residue correlations are negative (green). Correlations relating successive residues are highlighted with gray rectangles. Strips along the carbon dimension (dark gray) indicate the position of the H/N plane along the carbon frequency. (c) HNCA, HNCACB, and HNCACO strip comparison for E220, which suffers from severe degeneracy problems (total acquisition 8 days 16 h). The hNCAnH was recorded in 6 days 12 h.

In addition to its relatively large size, the protein is mainly α-helical[10] and subject to spectral overlap. Conventional strategies assigned 124 residues out of 239 (picked in HNCA). With the hNCAnH experiment, 151 signals were assigned, two mistakes were corrected, and 11 additional connectivities were unambiguously established. Ambiguities in strip comparisons often require multiple hypotheses (that need to be verified or corrected a posteriori) involving knowledge of the sequence, only to establish a connectivity. The hNCAnH identifies sequential residues without a need to know the amino acid sequence.

The main advantage of the new experiment is the relative ease and speed of assignment based on striking visual patterns. This is demonstrated in Figure 2, which shows an example of chain elongation in EntB. First, we choose the (H, N) coordinates of a given residue, e.g., those subsequently assigned to Glu 220. We then look at the H/N plane of the hNCAnH experiment, at the frequency of the corresponding α carbon, C^{α}_{220}. The sequential cross-peaks (with opposite phases) enable the identification of the (H, N) coordinates of the successor, Val 221 (a). A cross section along C at (H^{N}_{221}, N_{221}) reveals the frequency of C^{α}_{221}. The corresponding H/N plane is then visualized (b) and displays the correlations between (H^{N}_{221}, N_{221}) and (H^{N}_{222}, N_{222}). As can be judged from the number of peaks in each plane, a number of degeneracies exist along the carbon dimension. In this example, conventional strip comparison would not allow for a determination of successive residues (Figure 2c). The power of the technique presented here relies in its ability to identify which carbons belong to sequentially neighboring residues, as opposed to chance correspondence of their frequencies.

The hNCAnH experiment enables the direct sequential connection of $^{1}H-^{15}N$ HSQC or TROSY-HSQC cross peaks and delivers the $^{13}C^{\alpha}$ assignments in addition. It facilitates the assignment of crowded spectra and enables NMR studies of systems heretofore inaccessible. For large systems, it is a useful supplement to the existing pool of triple-resonance experiments. For any system size, it provides a trivial and visually simple assignment procedure. For smaller proteins, the hNCAnH can be used to entirely replace conventional backbone assignment experiments, such as HNCA, HNCACB, or HNCACO, and save time. The experiment promises spectacular results when nonuniform sampling is used to obtain ultrahigh resolution spectra.[11] Because long transfer delays are used in the experiment, high resolution can be achieved with little reduction in sensitivity. For very large systems, we have developed an alternative CANCA ^{13}C detected version.[12] The perdeuterated protein can then be dissolved in D_2O, so that all nuclei are subject to small relaxation rates, which further extends the limit in size of systems that can be assigned. This is however only beneficial when relaxation losses in the hNCAnH surpass the polarization advantages of proton detected experiments. Overall, the hNCAnH experiment provides a new perspective in studies of many systems such as large proteins, a number of α-helical proteins, and unfolded proteins.

Acknowledgment. This research was supported by the National Institutes of Health (Grants GM47467, AI37581, and RR00995) and a Human Frontier Science Program long-term fellowship from 2007-2009 (A.K.). A.K. is a Director's fellow of the Los Alamos National Laboratory awarded in 2009.

Supporting Information Available: This contains details of sample preparation, a description of the pulse sequence with alternative versions and further discussions. The pulse sequences can be downloaded at http://gwagner.med.harvard.edu. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Sattler, M.; Schleucher, J.; Griesinger, C. *Progr. NMR Spectrosc.* **1999**, *34*, 93–158. (b) Ferentz, A. E.; Wagner, G. *Q. Rev. Biophys.* **2000**, *33*, 29–65.
(2) Weisemann, R.; Ruterjans, H.; Bermel, W. *J. Biomol. NMR* **1993**, *3*, 113–20.
(3) Panchal, S. C.; Bhavesh, N. S.; Hosur, R. V. *J. Biomol. NMR* **2001**, *20*, 135–47.
(4) (a) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 4369. (b) Matsuo, H.; Kupce, E.; Li, H.; Wagner, G. *J. Magn. Reson. B* **1996**, *111*, 194–8. (c) Bracken, C.; Palmer, A. G., 3rd; Cavanagh, J. *J. Biomol. NMR* **1997**, *9*, 94–100.
(5) Sun, Z. Y.; Frueh, D. P.; Selenko, P.; Hoch, J. C.; Wagner, G. *J. Biomol. NMR* **2005**, *33*, 43–50.
(6) Frueh, D. P.; Arthanari, H.; Wagner, G. *J. Biomol. NMR* **2005**, *33*, 187–96.
(7) Frueh, D. P.; Sun, Z. Y.; Vosburg, D. A.; Walsh, C. T.; Hoch, J. C.; Wagner, G. *J. Am. Chem. Soc.* **2006**, *128*, 5757–5763.
(8) (a) Frueh, D. P.; Arthanari, H.; Koglin, A.; Vosburg, D. A.; Bennett, A. E.; Walsh, C. T.; Wagner, G. *Nature* **2008**, *454*, 903–6. (b) Koglin, A.; Lohr, F.; Bernhard, F.; Rogov, V. V.; Frueh, D. P.; Strieter, E. R.; Mofid, M. R.; Guntert, P.; Wagner, G.; Walsh, C. T.; Marahiel, M. A.; Dotsch, V. *Nature* **2008**, *454*, 907–11.
(9) Gehring, A. M.; Bradley, K. A.; Walsh, C. T. *Biochemistry* **1997**, *36*, 8495–503.
(10) Drake, E. J.; Nicolai, D. A.; Gulick, A. M. *Chem. Biol.* **2006**, *13*, 409–19.
(11) Rovnyak, D.; Frueh, D. P.; Sastry, M.; Sun, Z. Y.; Stern, A. S.; Hoch, J. C.; Wagner, G. *J. Magn. Reson.* **2004**, *170*, 15–21.
(12) Takeuchi, K.; Frueh, D.; Hyberts, S.; Sun, Z. Y.; Wagner, G. In preparation.

JA9046685