Journal of Biomolecular NMR, 3 (1993) 233–237
ESCOM

J-Bio NMR 112

A simple way for sequential assignment in isotopically enriched proteins using a H(N)CACO correlation

Stephan Seip, Jochen Balbach and Horst Kessler*

Organisch Chemisches Institut, TU München, Lichtenbergstraße 4, 8046 Garching, Germany

Received 17 November 1992
Accepted 26 January 1993

Keywords: Triple resonance; Sequential assignment; Isotope labeling; Mannose permease; Proteins

SUMMARY

A simplified scheme for sequential assignment in isotopically enriched proteins is presented. It is based on the standard triple resonance experiments HNCO, HN(CO)CA, HNCA and a modified H(N)CACO correlation, in which both of the H^N connectivities to the CA/C′ pair of residue i and i–1 are observed. The H(N)CACO was tested on uniformly $^{13}C/^{15}N$ enriched P13 domain of mannose permease (31 kDa).

Recent advances in multidimensional high-resolution NMR spectroscopy allow for the complete unambiguous assignment of isotopically enriched proteins (Ikura et al., 1990; Kay et al., 1990; Clore and Gronenborn, 1991). A crucial step is the assignment of the backbone atoms H^N, HA, CA, N, and C′. Sufficient correlations should, in principle, be present in the 'standard' HNCA, HNCO, HCACO, HCA(CO)N (Kay et al., 1990), HN(CO)CA (Bax and Ikura, 1991), and HN(CA)H spectra (Clubb et al., 1992a, Seip et al., 1992). In most cases, however, essential information is also obtained from conventional TOCSY-HMQC and NOESY-HMQC spectra. An alternative approach uses the information from just two 4D spectra for the complete backbone assignment (Kay et al., 1992). These methods appear to work well for proteins up to 20 kDa in size.

Serious problems occur for larger proteins (>30 kDa), when no homonuclear proton TOCSY transfer is possible. This, together with CA and HA chemical shift degeneracy, is the source of many problems during the sequential assignment (Greziek et al., 1992).

One way to resolve the ambiguities caused by chemical shift degeneracy is to make use of the intraresidual correlation to the carbonyl via HN(CA)CO, as proposed by Clubb et al. (1992b). This correlation is achieved using the CA nucleus for relay of magnetization from the NH to the carbonyl. Here we propose a similar pulse scheme H(N)CACO, which uses the nitrogen nucleus

* To whom correspondence should be addressed.

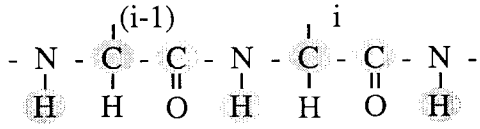

Fig. 1. Peptide backbone showing the nuclei correlated by the H(N)CACO.

for relay, instead of the CA (Fig. 1). This minor modification of the original experiment, however, has substantial consequences for the sequential assignment procedure.

The key step of the backbone sequential assignment is the correlation of an amino acid to its predecessor. This is established by inspection of HNCO and HN(CO)CA spectra which yield the CA_{i-1} and C'_{i-1} resonances. The HA_{i-1} proton that belongs to this pair of resonances can then be obtained from an HCACO spectrum and confirmed in the HN(CA)H. There is, however, no straightforward way to obtain the $^{15}N_{i-1}$ and H^N_{i-1} resonances that belong to the HA-CA-C′ triplet. The high degree of spectral overlap and small uncertainties in the peak positions in the different spectra usually result in ambiguous connectivities. The correct H^N_{i-1} can usually be identified with relative ease, from d_{NN} connectivities in NOESY spectra. In β-sheets or loop regions, however, this approach is of little use, because overlap in the HA region of the NOESY spectrum makes an unambiguous assignment of the H^N_{i-1} almost impossible. The H(N)CACO experiment can overcome most of the aforementioned problems. It will fail only in the unlikely case of degeneracy of CA-C′ pairs. It must be noted that the sensitivity of the experiment is considerably lower than that of the HNCO, HNCA or HN(CO)CA experiments. We believe, however, that the information present in the H(N)CACO spectrum justifies a prolonged measurement time (in our case, 4.5 days for a 31 kDa protein). Although the same information can in

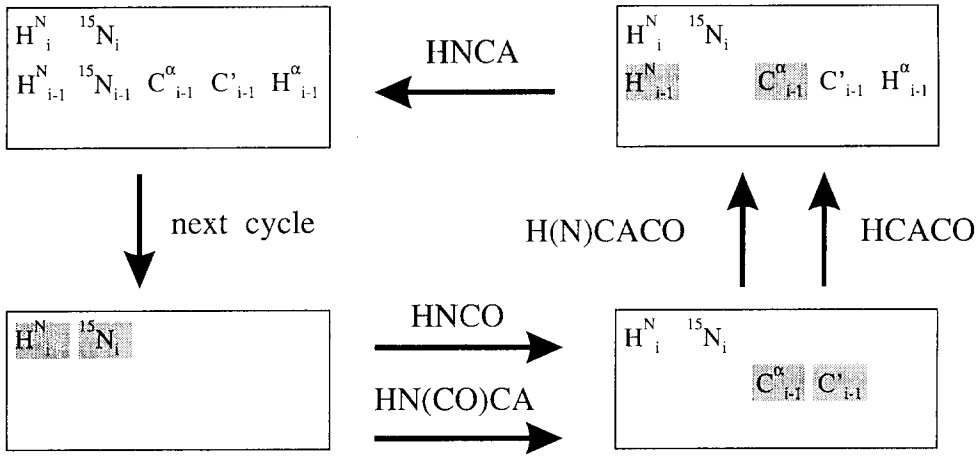

Fig. 2. Scheme for the sequential assignment in the H(N)CACO experiment. Each box represents the chemical shift list, for which the entries for residue (i–1) are completed during a cycle. The shifts indicated by a filled box are used as the *two* known parameters for the next search step. For example, both H^N and N chemical shifts are known when looking for CA_{i-1} and C'_{i-1}. Each search step adds only *one* new chemical shift (for each arrow) to the list, with two shifts fixed. The H(N)CACO serves for the intraresidue connection of the CA and C′ to the H^N. This avoids most problems associated with uncertainties in picked peak maxima from different spectra (HNCA and HN(CA)CO, HCACO).

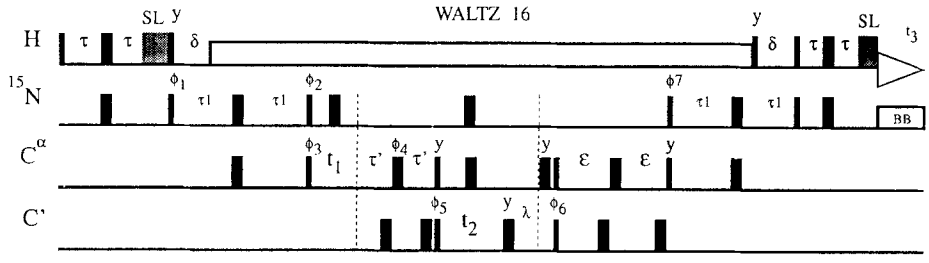

Fig. 3. Pulse sequence of the H(N)CACO experiment. Phase cycling: $\phi1$ = x, −x; $\phi2$ = x, x, −x,−x; $\phi3$ = x; $\phi4$ = 4x, 4(y), 4(−x), 4(−y); $\phi5$ = 32(y), 32(−y); $\phi6$ = 8(y), 8(−y); $\phi7$ = 16(x), 16(−x). Receiver: x,−x,−x,x, 2(−x,x,x,−x), x,−x,−x,x, 2(−x,x,x,−x, 2(x,−x,−x,x), −x,x,x,−x), x,−x,−x,x, 2(−x,x,x,−x), x,−x,−x,x. All other pulses are applied along the x-axis. Sign discrimination in F1 and F2 was achieved by TPPI of phases $\phi3$ and $\phi5$. A synchronous WALTZ16-(x) sequence was used for proton decoupling (field strength was reduced to 8.3 kHz during decoupling). Immediately after decoupling, a 90°(y) purge pulse was applied on protons. Broadband decoupling during acquisition was achieved with the GARP-2 sequence. τ = 2.25 ms, τl = 12.9 ms, δ = 4.5 ms. τ' is decremented with increasing t_1 from its initial value of 3.6 ms to allow for evolution of C^α chemical shift. Water suppression was achieved by weak presaturation during the relaxation delay and the two spin-lock pulses, SL. The carbon transmitter was centered in the C^α region and the carbon pulse length was set to 56 μs, to minimize excitation of the carbonyl resonances. Carbonyl 90°-pulses were applied as a series of phase shifted pulses of 9.2 μs duration with 60° phase shifts for off-resonance excitation.

principle be obtained by inspection of both HNCA and HN(CA)CO spectra, it is in practice much easier to use the H(N)CACO spectrum. Figure 2 shows schematically the search procedure that the H(N)CACO utilizes. It is apparent that the assignment is straightforward and less prone to errors caused by small deviations of chemical shifts in the different spectra. The main advantage of this method is that in all of the steps of the cycle only *one* new chemical shift must be found in the corresponding spectrum.

The pulse sequence of the H(N)CACO experiment is given in Fig. 3. The details of this type of correlation have been reported elsewhere (Clubb et al., 1992b), and we will give only a brief description of the differences with respect to the original sequence. The initial INEPT sequence transfers magnetization from ^1H to ^{15}N. During the delay, 2τl, antiphase magnetization of ^{15}N with respect to CA is created with simultaneous refocusing of H–N antiphase magnetization and synchronous proton decoupling. In the following INEPT transfer to C′, chemical shifts of CA evolve during t_1 in a constant time period of $2\varepsilon = t_1 + 2\tau'$. To compensate for Bloch–Siegert phase errors during the evolution of CA magnetization, caused by the C′ decoupling pulse (Vuister and Bax, 1992), an additional 180° pulse on C′ is placed immediately before the first 90° pulse on C′. To apply the same procedure to C′, a refocusing pulse is used during t_2, to reverse the precession of C′ magnetization. A zero-phase correction for this indirect dimension can then be easily obtained by setting the delay λ to $2t_2(0)$ (Schmieder et al., 1991). Magnetization is finally transferred back to H^N by three successive INEPT transfers and detected during t_3.

We applied the sequence to a 1 mM (dimer) sample of uniformly $^{15}N/^{13}C$-enriched P13 domain of mannose permease of *E. coli* (Erni et al., 1987) in H_2O at 37 °C. P13 consists of 135 residues and has a total weight of 31 kDa as uniformly ^{13}C- and ^{15}N-labeled homodimer. The experiment was recorded on an unmodified Bruker AMX-600 at 37 °C. 64 × 96 × 1024 real points were taken in F1, F2, and F3, respectively. Acquisition times were 7 ms (F1), 23 ms (F2), and 122 ms (F3).

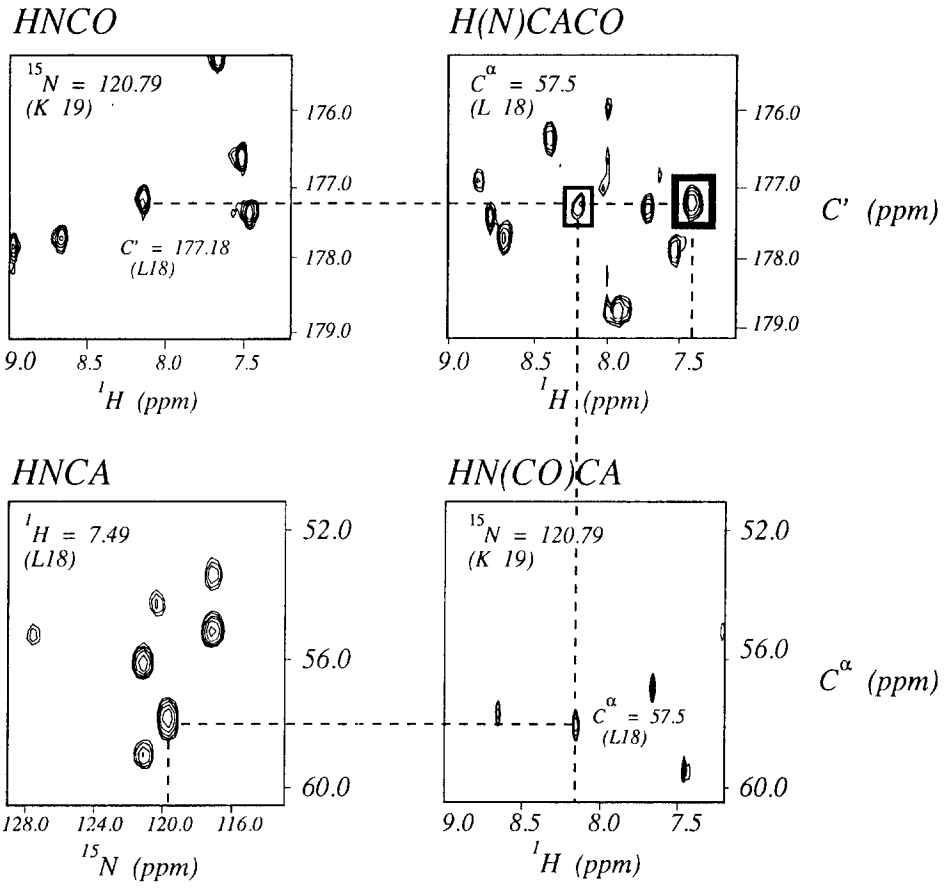

Fig. 4. Sequential assignment of the K19 and L18 residues in P13 using the H(N)CACO. The first step was to find the CA_{i-1} and C'_{i-1} resonances of L18 (at 57.5 and 177.2 ppm, respectively) in the HN(CO)CA and HNCO spectra. The resonance of H^N_{i-1} (7.49 ppm, L18) was then immediatedly found at the coordinates of CA_{i-1} and C'_{i-1} (thick box). The nitrogen resonance of L18 was eventually found in the HNCA. The presence of the sequential peak to the H^N_i (8.26 ppm, thin box) of K19 greatly facilitated the assignment.

Sixty-four scans were taken for each increment, resulting in a total measurement time of 109 h. The spectrum consists of $64 \times 128 \times 512$ (real) data points. A 70°-shifted squared sinebell was used for apodization in all dimensions. Figure 4 shows cross sections through HN(CO)CA, HNCO, HNCA, and H(N)CACO spectra used for the sequential assignment of residues K19 and L18 of P13. 105 of the 129 expected interresidual connectivities and 67 of the 128 possible intraresidual connectivities were obtained from the H(N)CACO.

In conclusion, we have shown that the H(N)CACO experiment can form an important role in a simplified and robust assignment strategy for backbone resonances of isotopically enriched proteins in the 20–30 kDa range. Our study of P13 shows, however, that the application of the H(N)CACO experiment is limited and that its sensitivity will not be sufficient for proteins larger than approximately 30 kDa.

ACKNOWLEDGEMENTS

We thank K. Flückiger and Prof. Dr. B. Erni for providing the sample of P13 and Dr. D. Hare for the Felix Program. This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

REFERENCES

Bax, A. and Ikura, M. (1991) *J. Biomol. NMR*, **1**, 99–104.
Clore, G.M. and Gronenborn, A.M. (1991) *Prog. Nucl. Magn. Reson. Spectr.*, **23**, 43–92.
Clubb, R.T., Thanabal, V. and Wagner, G. (1992a) *J. Biomol. NMR*, **2**, 203–210.
Clubb, R.T., Thanabal, V. and Wagner, G. (1992b) *J. Magn. Reson.*, **97**, 213–217.
Erni, B., Zanolari, P. and Kocher, H.P. (1987) *J. Biol. Chem.*, **262**, 5238–5247.
Greziek, S., Döbeli, H., Gentz, R., Garotta, G., Labhardt, A.M. and Bax, A. (1992) *Biochemistry*, **31**, 8180–8190.
Ikura, M., Kay, L.E. and Bax, A. (1990) *Biochemistry*, **29**, 4659–4667.
Kay, L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **89**, 496–514.
Kay, L.E., Wittekind, M., McCoy, M.A., Friedrichs, M.S. and Mueller, L. (1992) *J. Magn. Reson.*, **98**, 443–450.
Schmieder, P., Zimmer, S. and Kessler, H. (1991) *Magn. Reson. Chem.*, **29**, 375–380.
Seip, S., Balbach, J. and Kessler, H. (1992) *J. Magn. Reson.*, **100**, 406–410.
Vuister, G.W. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 428–435.