Unambiguous Resonance Assignments in $^{13}C,^{15}N$-Labeled Nucleic Acids by 3D Triple-Resonance NMR

Bennett T. Farmer II* and Luciano Müller

Macromolecular NMR, Pharmaceutical Research Institute
Bristol-Myers Squibb, P.O. Box 4000
Princeton, New Jersey 08543-4000

Edward P. Nikonowicz and Arthur Pardi

Department of Chemistry and Biochemistry
University of Colorado at Boulder
Boulder, Colorado 80309-0215

Received June 7, 1993

Multidimensional heteronuclear magnetic resonance experiments are revolutionizing the solution-structure determinations of isotopically labeled proteins and nucleic acids.[1,2] One recent advance in this area has been the unambiguous sequential assignment of amino acid residues in $^{13}C,^{15}N$ isotopically labeled proteins using 3D and 4D $^{1}H,^{13}C,^{15}N$ triple-resonance NMR correlation experiments, which rely solely on through-bond scalar connectivities.[3-5] Current assignment techniques in nucleic acids, however, still employ conformation dependent, through-space NOE interactions to establish intraresidue correlations between the sugar and base protons. Such techniques suffer from the inability to *unambiguously* distinguish interresidue from intraresidue interactions. To alleviate this ambiguity, we present here the first application of a through-bond, 3D $^{1}H,^{13}C,^{15}N$ triple-resonance NMR correlation experiment to the assignment of $^{13}C,^{15}N$ isotopically labeled RNAs.

Figure 1 depicts the $H_sC_s(NC)_bH_b$ 3D $^{1}H,^{13}C,^{15}N$ triple-resonance experiment used to unambiguously correlate the H1′ and C1′ assignments of the ribose to the H6/8 assignment of the base within each nucleotide. The subscript s in the sequence name denotes spins on the ribose; and b, spins on the base. The $H_sC_s(NC)_bH_b$ experiment seeks to provide H1′(t_3)–C1′(t_2)–(N1/9–C6/8)–H6/8(t_1) correlations in nucleotides. The pulse sequence is more complex than its protein-applied counterparts due to the more intricate scalar-coupling network within the nucleotide base (*vide infra*). The delay, $\tau_{CN} = \tau_4 + \tau_5 + \tau_6$ (Figure 1), during which time the N1/9 coherence evolves to become antiphase to C6/8, respectively, has initially been optimized at 30 ms for guanosine. For pyrimidines, this τ_{CN} value incurs an ~10% loss in sensitivity compared to their optimum τ_{CN} value of ~42 ms. The difference in optimum τ_{CN} values is primarily due to the relatively large N9–N3 scalar coupling in guanosine, which cannot be selectively refocused due to chemical shift overlap between the N9 and N3 resonances in this nucleotide.[10] Although the same N9–N3 coupling exists for adenosine, one can selectively refocus it since the chemical shift of its N3 is approximately 47 ppm downfield of N9.[10]

Figure 2 presents the $^{1}H,^{13}C,^{1}H$ intraresidue sugar-to-base correlations for C3, U6, U11, and C12 in the 99% $^{13}C,^{15}N$-labeled RNA duplex r(GGCGCUUGCGUC)$_2$.[13] Note that the extremely high resolution obtained in the C1′ dimension allows for the separation of closely spaced carbon resonances. The sugar-to-base correlations in Figure 2 are anchored to the H1′,C1′ resonance pair, requiring only that no two pyrimidines have the same H1′ *and* C1′ chemical shifts. From these data, the H6 chemical shift of each pyrimidine base can be correlated to the H1′ and C1′ chemical shifts of its respective ribose. Unambiguous intraresidue sugar-to-base correlations have been observed for all seven pyrimidines in this duplex and are consistent with previous assignments.[14] The sensitivity of the $H_sC_s(NC)_bH_b$ experiment is also good considering the overall length of the sequence. U7 yields the weakest correlation, potentially due to the presence of a minor conformer for this nucleotide (*unpublished observations*). U6 and U11 present the next weakest correlations, being only 45% more intense than that for U7. The most intense correlation

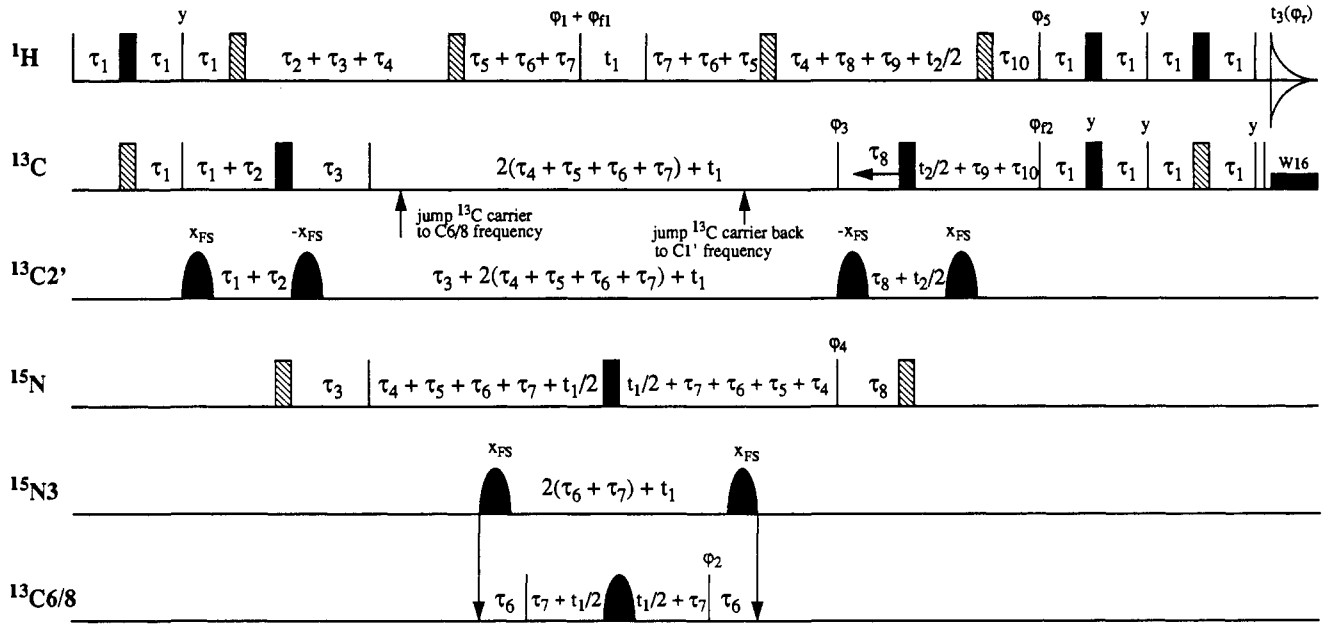

Figure 1. The 3D $^{1}H,^{13}C,^{15}N$ triple-resonance sequence $H_sC_s(NC)_bH_b$ used to establish intraresidue sugar-to-base correlations in $^{13}C,^{15}N$-labeled RNA pyrimidine nucleotides. 90° pulses are represented by narrow lines; simple 180° pulses, by black rectangles; and composite $90_x240_y90_x$ inversion pulses,[6] by diagonally striped rectangles. Unless otherwise indicated, all pulses have phase x. The phase cycle is $\varphi_1 = x, -x$; $\varphi_2 = 2(x), 2(-x)$; $\varphi_3 = 8(x), 8(-x)$; $\varphi_4 = 4(x), 4(-x)$; $\varphi_5 = 16(x), 16(-x)$; and $\varphi_r = \varphi_1 + \varphi_2 + \varphi_3 + \varphi_4 + \varphi_5$. Complex data were collected in t_1 by States-TPPI[7] with FIDs for $\varphi_{f1} = x, y$ being stored separately. Complex data were collected in t_2 with FIDs for $\varphi_{f2} = (x, -x)^8$ being stored separately. The 3D hypercomplex data set was processed as described by Palmer *et al.*[8] to achieve a quadrature-detected, phase-sensitive display along the C1′ dimension with sensitivity enhancement. *All* pulses on lines labeled with a particular *spin group* were Gaussian-shaped (64 steps, 5σ cutoff) and are selective for that spin group. The subscript FS appended to certain pulse phases denotes a frequency-shifted pulse achieved by phase modulation.[9]

0002-7863/93/1515-11040$04.00/0

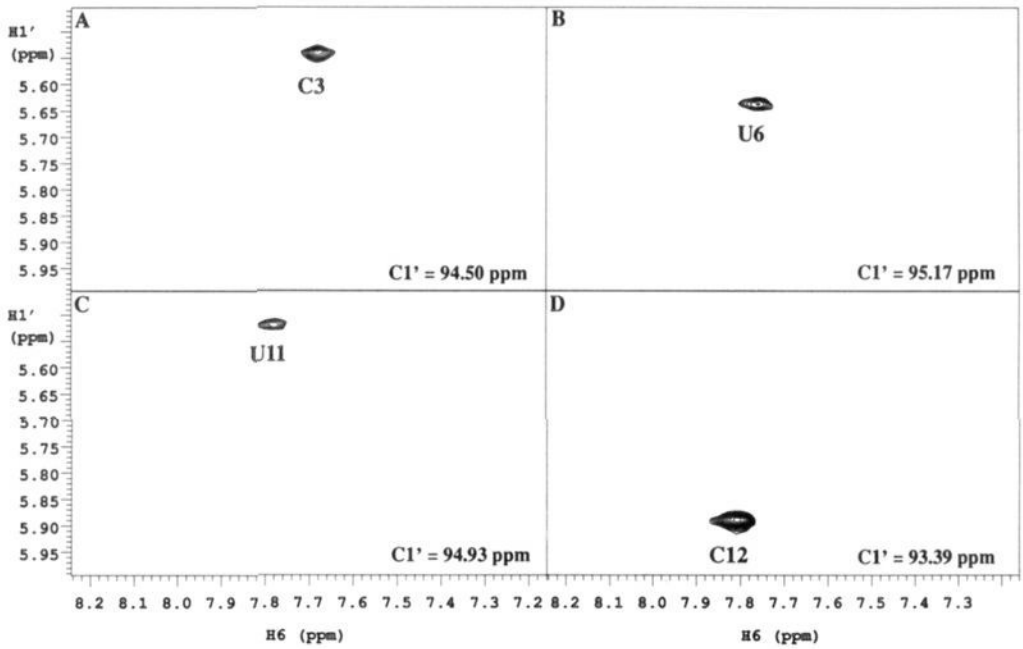

Figure 2. 2D planes from the $H_sC_s(NC)_bH_b$ experiment used to complete the intraresidue sugar-to-base correlation for (A) C3; (B) U6; (C) U11; and (D) C12. Each plane is taken at the respective C1′ chemical shift for that nucleotide. All spectra were recorded on a UNITY 600 spectrometer at 30 °C with $t_{90}(^1H) = 7.5\ \mu s$, $t_{90}(^{13}C) = 13.4\ \mu s$, $t_{180}(C2') = 896\ \mu s$, $t_{90}(C6) = t_{180}(C6) = 704\ \mu s$, γB_2 (C1′ decouple) = 1.29 kHz with WALTZ16,[11] $t_{90}(^{15}N) = 37\ \mu s$, sw(H6) = 650.0 Hz, t_1^{max}(H6) = 35.38 ms, sw(C1′) = 600.0 Hz, t_2^{max}(C1′) = 25.00 ms, sw(^{1}H) = 3 kHz, t_3(H1′) = 120.0 ms, τ_1 = 1.48 ms, τ_2 = 18.52 ms, τ_3, = 20.00 ms, τ_4 = 16.20 ms, $\tau_5 = t_1/4$, τ_7 = 2.40 ms, $\tau_6 = \tau_4 - \tau_5 - \tau_7$, $\tau_8 = \tau_3 - (t_2/2)$, $\tau_9 = \tau_2 - (t_2/2)$, and $\tau_{10} = \tau_1 + (t_2/2)$. The ^{13}C carrier was placed at the center of the C1′ resonances except where indicated in Figure 1. The C2′ selective pulses were shifted by –2550 Hz from the ^{13}C carrier frequency.[9] No N3 selective inversion pulses were applied due to the absence of adenosine in this RNA sample. The total acquisition time was 26 h. Linear prediction was used to extend the time-domain data in t_1 and t_2 prior to Fourier transformation.[12] The final data size was 128 × 128 × 1024 real points ($F_1F_2F_3$).

is observed for C12, probably because of more favorable relaxation rates in this terminal nucleotide. The intensity of the correlation for C12 is approximately 3–4 times that for the other six pyrimidines.

Because only one guanosine out of five in this duplex yielded any measurable signal in the $H_sC_s(NC)_bH_b$ experiment, we have devised a different assignment strategy for purines, mainly guanosine, which will be published elsewhere.[15] The RNA duplex in this study contains no adenosines. $H_sC_s(NC)_bH_b$ experiments on ^{13}C,^{15}N-labeled mononucleotides, however, indicate that AMP presents more favorable results than GMP (*data not shown*). The $H_sC_s(NC)_bH_b$ triple-resonance method for establishing intraresidue sugar-to-base correlations should therefore accommodate adenosine more readily than guanosine. Future work will be aimed at extending this class of triple-resonance experiments to larger isotopically labeled RNAs, e.g., ribozymes and tRNA.

Acknowledgment. This work was supported by NIH Grant AI33098 and NIH Research Career Development Award AI01051 to A.P. and by an NIH NRSA fellowship to E.P.N. We also thank the W. M. Keck Foundation for its generous support of RNA science on the Boulder campus.

* To whom correspondence should be sent.

(1) Clore, G. M.; Gronenborn, A. M. *Science* **1991**, *252*, 1390–1399.
(2) Nikonowicz, E. P.; Pardi, A. *Nature* **1992**, *355*, 184–186.
(3) Ikura, M.; Kay, L. E.; Bax, A. *Biochemistry* **1990**, *29* (19), 2659–2667.
(4) Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1990**, *89*, 496–514.
(5) Olejniczak, E. T.; Xu, R. X.; Petros, A. M.; Fesik, S. W. *J. Magn. Reson.* **1992**, *100*, 444–450.
(6) Levitt, M. H. *Prog. NMR Spectrosc.* **1986**, *18*, 61–122.
(7) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393–399.
(8) Palmer, A. G., II; Cavanagh, J.; Wright, P. E.; Rance, M. *J. Magn. Reson.* **1991**, *93*, 151–170.
(9) Patt, S. L. *J. Magn. Reson.* **1992**, *96*, 94–102.
(10) Büchner, P.; Maurer, W.; Rüterjans, H. *J. Magn. Reson.* **1978**, *29*, 45–63.
(11) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335–338.
(12) Zhu, G.; Bax, A. *J. Magn. Reson.* **1990**, *90*, 405–410.
(13) The sample is 0.9 mM in duplex (1.8 mM single strand) with 10 mM K_2PO_4, 150 mM NaCl, and 0.02 mM EDTA at pH 6.8 in D_2O.
(14) Nikonowicz, E. P.; Pardi, A. *J. Mol. Biol.* **1993**, *232*, 1141–1156.
(15) Farmer, B. T., II; Müller, L.; Nikonowicz, E.; Pardi, A. *J. Biomol. NMR*, submitted for publication.